Investigating the functional and structural modulation of carboxysomes in *Synechococcus elongatus* PCC7942

Thesis submitted in accordance with the requirements of the

University of Liverpool for the degree of Doctor in Philosophy by

Yaqi Sun

September 2018

Department of Functional and Comparative Genomics, School of Biological Science, Institute of Integrative Biology

Abstract

Organelle formation and compartmentalisation within eukaryotic and prokaryotic cells provide the structural foundation for modulation of metabolic reactions in space and time. Bacterial microcompartments (BMCs) are self-assembling organelles widespread among bacterial phyla. By physically sequestering specific enzymes key for metabolic processes from the cytosol, these organelles play essential roles in carbon fixation, microbial ecology and pathogenesis. Carboxysomes serve as the key CO_2-fixing machinery in all cyanobacteria and some chemoautotrophs. The β-carboxysomes in the cyanobacterium *Synechococcus elongatus* PCC7942 (Syn7942) have been extensively characterised as the model carboxysomes. In **Chapter 1**, general research background centralised on carboxysomes was summarised. In **Chapter 2**, material and methods adopted in this thesis were documented. From **Chapter 3** to **Chapter 5**, I addressed specific topics on functional and structural modulation of carboxysomes in Syn7942. In **Chapter 6**, I concluded the findings in previous chapters and showed the perspectives following this project.

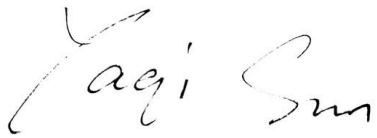
In **Chapter 3**, we used fluorescence tagging and live-cell confocal fluorescence imaging to explore the biosynthesis and subcellular localisation of β-carboxysomes within Syn7942 cell in response to light variation. We demonstrated that β-carboxysome synthesis on cellular level is re-modulated in response to increasing light intensity, thereby enhancing the carbon fixation activity of the cell. Inhibition of photosynthetic electron flow impairs the accumulation of carboxysomes, indicating close coordination between β-carboxysome biogenesis and photosynthetic electron transport. Likewise, the spatial organisation of carboxysomes in the cell correlates with the redox state of photosynthetic electron transport chain. In **Chapter 4**, we used live-cell single-molecule fluorescence microscopy, coupled with confocal and electron microscopy, to decipher the absolute protein stoichiometry and organisational variability of single β-carboxysomes in Syn7942 cell. I find that the protein stoichiometry, diameter, cellular localisation and mobility pattern of carboxysomes in cells depend sensitively on the microenvironmental levels of CO_2 and light intensity during cell growth, revealing cellular strategies of dynamic regulation. In **Chapter 5**, I documented the biosynthesis and organisation of carboxysomes under diurnal dark-light cycles compare to constant light in Syn7942. We found the reoccurring carboxysome distribution changes and rhythmic carbon fixation capacities in diurnal condition. Also, we documented the localisation, enzymatic activity as well as the quantity of carboxysomes in circadian null background to provide preliminary evidence of circadian control in carboxysome biogenesis.

The findings in this thesis provide essential knowledge for us to modulate the β-carboxysome biosynthesis and function in cyanobacteria. Furthermore, improving our understanding of carboxysome assembly principles will aid rational design of functional metabolic factories in heterologous organisms for metabolic engineering using synthetic biology.

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Liverpool or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Liverpool or any other University or similar institution except as declared in the Preface and specified in the text. This thesis does not exceed the prescribed word limit of 100,000 words.

Yaqi Sun



Acknowledgement

To start with, I would like to express my greatest gratitude to my supervisor Dr Luning Liu for his support and guidance during these years. Luning has been a great tutor in the field of science, and a supportive friend in life, especially during my difficulties. His determination and dedication to the research have inspired me to commit more to the research work, providing me with a role model to follow in the future.

I would like to acknowledge the support from China Scholarship Council that gives me the opportunity to extend my science knowledge in a wonderful university. Furthermore, express my deep gratitude to the care and support provided by the Education Section, the Consulate General of the people's republic of China in Manchester.

I would like to thank all my colleagues, Selene Casella, Yi Fang, Fang Huang, Taiyu Chen, Mengru Yang, Zimeng Zhang, Qiuyao Jiang, Longsheng Zhao, Tianpei Li, Matthew Faulkner, Greg Dykes, Leanne Miller, Jorge Rodriguez-Ramos, Tuomas Huokko and Monsour Hazeem in the Liu Lab who shared laughter and tears along the road. I will always remember this wonderful big family.

I would like to thank everyone in the Lab G and the institute especially our technicians, Jean Wood and Paul Loughane for their assistance in daily research and Dr Meriel Jones for the kind help and suggestions during my PhD study. I would like to thank all the staff in CCI centre, especially Dr David Mason, Dr Marco Marcello and Ms Jennifer Adcott and Ms Joanna Wnetrzak for their kind help with the use of fluorescent microscopy. Their kindness makes studying at Liverpool a delightful experience that I would always remember.

Many thanks to our collaborators, Prof Mark Leake and Dr Adam Wollman at the University of York, Dr Steve Barrette at Physics department in the University of Liverpool, Professor Dean Price and Dr Benedict Long at Australian National University. Without your assistance, none of the work in this thesis could be achieved.

I would like to thank my wife who sacrifices her own career to stay with me during my PhD study. Home is always with me with her around.

In memory of my dear mom.

Publications and Author's contributions

The major content of **Chapter 3** was published in article entitled: Light modulates the biosynthesis and organisation of cyanobacterial carbon fixation machinery through photosynthetic electron flow. Sun, Y., Casella, S., Fang, Y., Huang, F., Faulkner, M., Barrett, S., & Liu, L. -N. (2016). Light Modulates the Biosynthesis and Organization of Cyanobacterial Carbon Fixation Machinery through Photosynthetic Electron Flow. *Plant Physiology*, 171(1), 530-541. doi:10.1104/pp.16.00107.

The major content of **Chapter 4** was under revision for The Plant Cell entitled: Single-organelle quantification reveals the stoichiometric and structural variability of carboxysomes dependent on the environment. Yaqi Sun, Adam J. M. Wollman, Fang Huang, Mark C. Leake, & Lu-Ning Liu.

I acknowledge the following collaborators for their contributions over the outcome described in the following chapters. Unless specified below, all works were completed by the author.

Name and organisation	Contributions	Related chapters
Prof Luning Liu University of Liverpool, UK	Guidance of research, revision of the manuscript	All chapters
Selene Casella University of Liverpool, UK	TEM imaging of whole cell Syn7942	Chapter 3
Yi Fang University of Liverpool, UK	Construction of GFP free vector	Chapter 3
Dr Fang Huang University of Liverpool, UK	Induction and imaging of GFP free strain; Construction of CcmO-eYFP:: RbcL-CFP strain	Chapter 3, 4
Matthew Faulkner University of Liverpool, UK	Transformation, maintenance of CcmK4:GFP strain	Chapter 3
Dr Steve Barrett University of Liverpool, UK	Development of microcompartment analysis plugin in software Image SXM	Chapter 3
Prof Mark Leake, University of York, UK	Slimfield data integration and revision of the manuscript	Chapter 4
Dr Adam Wollman, University of York, UK	Slimfield imaging and analysis	Chapter 4

Other publications produced during PhD studies:

Roles of RbcX in Carboxysome Biosynthesis in the Cyanobacterium Synechococcus elongatus PCC7942 Fang Huang, Olga Vasieva, Yaqi Sun, Matthew Faulkner, Gregory F. Dykes, Ziyu Zhao, Lu-Ning Liu. *Plant Physiology* 179 (1) 184-194; DOI: 10.1104/pp.18.01217.

I performed the immunoblot detection of RbcX-YFP.

List of Contents

Preface

List of Figures

Preface

List of Tables

Alphabetical list of Abbreviations and Acronyms

PGA (3PGA)	3-phosphoglycerate	HS-AFM	High-speed Atomic Force Microscopy
AFM	Atomic Force Microscopy	IPTG	Isopropyl-ß-D-thiogalactopyranoside
Chla	chlorophyll a	kb	pair
ADP	Adenosine diphosphate	kDa	kilo Dalton
AFM	atomic force microscopy	K_m	Michaelis constant
ATP	Adenosine triphosphate	KO	Knock-out
ATPase	ATP synthase	kV	kilo Volt
BMC(s)	Bacterial microcompartment(s)	LB	Luria-Bertani medium
bp	Base pair	LL	Low light
C6	Cytochrome C6	LSUs (SSUs)	Large subunits (small subunits)
CA	Carbonic anhydrase	ML	Moderate light
CAM	Crassulacean Acid Metabolism	MS	Mass spectroscopy
CB(s)	Carboxysome(s)	MSD	Mean square displacement
CBB cycle	Calvin-Benson-Bassham cycle	MTA	Mixed alkyltrimethylammonium bromide
CCM	Carbon concentrating mechanism	NADP+	Oxidised β-nicotinamide adenine dinucleotide phosphate
CFP	Cyan Fluorescent Protein	NADPH	Reduced β-nicotinamide adenine dinucleotide phosphate
Chl	chlorophyll	NDH	NADPH-dehydrogenase like complex
Ci	inorganic carbon	OD	Optical Density
CL	Constant Light	PAGE	Polyacrylamide Gel Electrophoresis
Cryo-EM	Cryogenic electron microscopy	PC	Plastocyanin
Cyt	Cytochrome	PCR	Polymerase chain reaction
DAPI	4’,6-diamidino-2-phenylindole	PDB	Protein database (https://www.rcsb.org/)
DBMIB	2,5- dibromo-3-methyl-6-isopropyl-benzoquinone	Pdu	Propanediol utilisation compartment
DCMU	3-(3,4-dichlorophenyl)-1,1-dimethylurea	P_i	Phosphate
DMSO	dimethyl sulfoxide	PQ	Plastoquinone
DNA	2’-deoxyribonucleic acid	PSI(II)	Photosystem I(II)
E. coli	*Escherichia coli*	PTO	Post-translational oscillator
eGFP (GFP)	enhanced Green Fluorescent Protein	R5P	Ribose-5-phosphate
EM	Electron microscopy	RPM	revolution per minute
Eut	Ethanolamine utilisation compartment	Ru5P	Ribulose-5-Phosphate
eYFP	enhanced Yellow Fluorescent Protein	Rubisco	Ribulose-1,5-bisphosphate carboxylase/oxygenase
Fd	Ferredoxin	RuBP	Ribulose-1,5-bisphosphate
FNR	Ferredoxin-NADP+ reductase	SD	Standard deviation
FP	Fluorescence protein	SDS	Sodium-Dodecyl-Sulfate
FRAP	Fluorescence recovery after photobleaching	Syn7942	*Synechococcus elongatus* PCC 7942
G3P	Glyceraldehyde 3-phosphate	WT	Wild-type
HCR	High-CO_2 Requirement	α-CcmK2	anti-CcmK2 antibody
HL	Higher light	α-RbcL	anti-RbcL antibody

Chapter 1

General introduction

Chapter 1

1.1 Bacterial Microcompartments

1.1.1 The compartmentalisation of bacterial cells

Organelle compartmentalisation within cells provides the structural foundation for physiological optimisation and modulation of metabolic reactions in space and time (Bobik et al., 2015; Kerfeld et al., 2010). In eukaryotic cells, the highly specialised subcellular organelles (such as chloroplasts, mitochondria and lysosomes) that encapsulate specific enzymes and metabolic pathways are well-separated from the cytosolic environment by controlled influx and efflux system for metabolites (Satori et al., 2013). By compartmentalisation, eukaryotic cells are capable of housing numerous biological pathways in the crowded cytosolic environment without mutual interference (Gabaldón and Pittis, 2015). Likewise, similar compartmentalisation is achieved through the self-assembling analogous organelles called bacterial microcompartments (BMCs) widespread among bacterial phyla (Axen et al., 2014). The BMCs structurally consist of polyhedral protein shells and interior enzymes which catalyse sequential metabolic reactions within lumens separated from the cytoplasm (Bobik, 2006; Kerfeld et al., 2010; Yeates et al., 2011). The BMC compartmentalisation enables diverse enzymatic reactions catalysing within the subcellular "micro-factories" at high efficiency, contributing to the metabolic diversities of bacteria accommodating specific habitats. Since the first BMC has been discovered back in 1956 (Drews and Niklowitz, 1956), tremendous research on BMC structure, assembly and function have been done in the past 60 years. Significant progress is achieved in the recent decades and reviewed systematically in the following review articles (Bobik, 2006; Bobik et al., 2015; Kerfeld et al., 2018; Kerfeld et al., 2010; Yeates et al., 2008).

1.1.2 Common features of BMC shells

Unlike organelles in eukaryotic cells that establish the border by lipid membrane systems in diverse forms, the boundaries of BMCs are formed by homologous icosahedral shells that self-assembled by thousands of proteins that belong to multiple protein paralogs (Yeates et al., 2010). The shells are single-layered (Kerfeld et al., 2005; Tanaka et al., 2008; Tsai et al., 2007), and structurally resembling the virus capsids when visualised under microscopy (Faulkner et al., 2017; Kerfeld and Erbilgin,

2015). The shells are consist of three conserved protein paralogs: the BMC-H domain family, the BMC-T domain family and BMC-P domain family proteins (Kerfeld and Erbilgin, 2015). The BMC-H proteins self-assemble into functional hexameric units and form the majority of the shell facets (Dryden et al., 2009; Kerfeld et al., 2010). The BMC-T proteins contain a tandem BMC-H domain which form up pseudo-hexamers (dimers of trimer) that make up a smaller fraction of the shells are suggested to be the potential gateways for shell permeability control with the open/close central pore through conformation changes (Cai et al., 2013; Klein et al., 2009). The BMC-P proteins are suggested to form the vertices of the shells in pentameric assembly units (Iancu et al., 2007; Sutter et al., 2013; Tanaka et al., 2008; Wheatley et al., 2013). The diagram of BMC shell formation is illustrated in **Figure 1-1**.

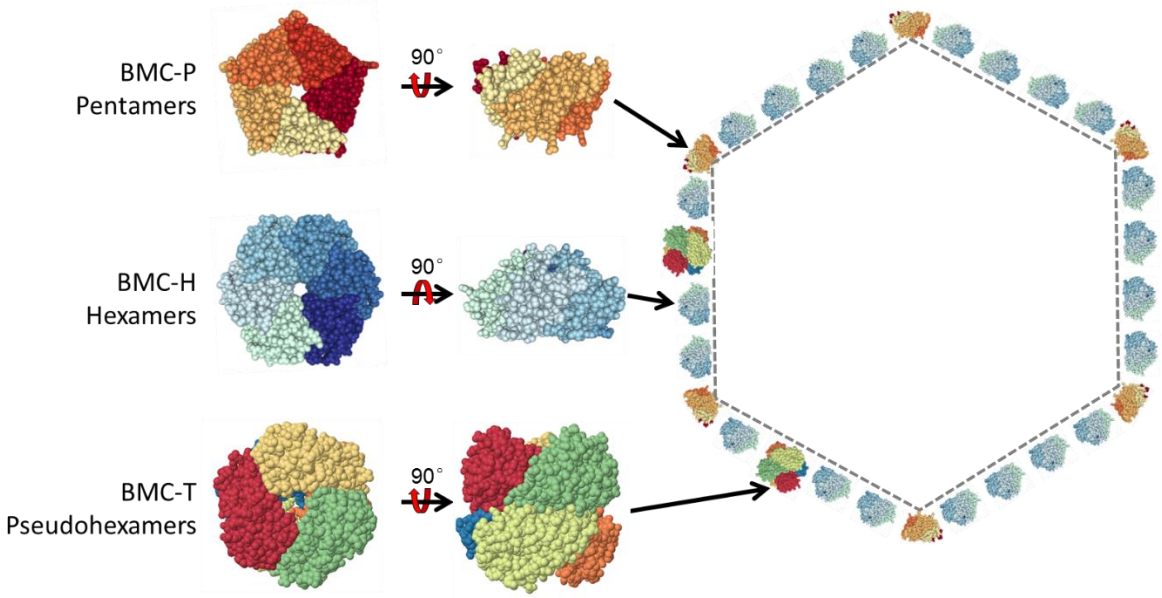


Figure 1-1. Schematic diagram of BMC shell assembly by pentameric and (pseudo)hexameric proteins. BMC-P proteins have been deduced to locate on the vertexes of the polyhedron shells, while BMC-H and BMC-T form the facets surfaces. Space fill model adopted from PDB entries (BMC-T PDB ID: 5LT5, BMC-H PDB ID: 2A1B; BMC-P PDB ID: 2QW7) reported in (Kerfeld et al., 2005; Sutter et al., 2017; Tanaka et al., 2008). Assembly was shown for illustration only; numbers and ratios of building blocks do not represent real stoichiometry.

BMC-P, BMC-H and BMC-T proteins have pronounced orientations with concave and convex surfaces (Kerfeld et al., 2005; Sutter et al., 2017; Tanaka et al., 2008). Ordered orientations of shell

proteins were confirmed through an artificial 6.5-MDa BMC shell assembly detected by cryogenic electron microscopy (cryo-EM) (Sutter et al., 2017). The well-defined shell structures operate as a physical barrier that concentrates and protect enzymes within the lumens (Fan et al., 2012). The distinctive properties of concave/convex of shell protein grant carboxysome the selective permeability to of substrates and products entering/exiting the shell pores (Chowdhury et al., 2015; Crowley et al., 2010; Park et al., 2017).

1.1.3 Diversity of BMCs

Although BMCs shells are structurally alike and phylogenetically related, a great variety of encapsulated enzymes has been found in bacteria survive in different environmental niches, fulfilling diverse biological functions (Kerfeld et al., 2018). Before the beginning of the genomic era, three types of BMCs were identified according to the biological processes housed inside the lumens: the carboxysome (commonly referred as CB) for CO_2 fixation, the metabolosome Pdu for Vitamin B_{12} dependent 1,2-propanediol utilisation, and the metabolosome Eut for ethanolamine utilisation. The anabolic carboxysomes are found in phototrophic cyanobacteria and some chemoautotrophic organisms (Price et al., 1998; Shively et al., 1973). The catabolic metabolosomes Pdu and Eut are discovered in bacteria with specialised heterotrophic metabolic system utilising organic carbons (Bobik et al., 1999; Kofoid et al., 1999). The distinctive biological processes within BMC lumens are fulfilled by the corresponding set of enzymes which are essential for CO_2 fixation, pathogenesis, and microbial ecology (Bobik et al., 2015; Yeates et al., 2010; Yeates et al., 2008). In the lumen, core enzymes that catalyse sequential metabolic reactions are packed underneath the shells in close proximity. The metabolic efficiencies are enhanced through concentrating substrates or avoiding leakage or toxic intermediates that are to be consumed sequentially (Chowdhury et al., 2014; Huang et al., 2001). Through encapsulation of enzymes, the unpreferable flux of reversible reactions is negated, meanwhile preventing toxicity damages caused by intermediates (if any) in the cytoplasm (Havemann et al., 2002; Yeates et al., 2008).

Up to date, comprehensive bioinformatic studies have indicated uncharacterised BMCs with a variety of functions widely spread among bacteria phyla (Axen et al., 2014). The functional study of these newly discovered BMCs has just initiated. For instance, the Planctomycetes and Verrucomicorbia microcompartment (PVM BMC) that capable of processing L-fucose and L-rhamnose (Erbilgin et al., 2014), the sub-type of Glycyl radical enzyme microcompartment (GRE BMC) with putative Vitamin B_{12} independent 1, 2-propanediol utilisation capacity (Zarzycki et al., 2017) have just been discovered. Meanwhile, with the progressively growing bacteria genomic database, more uncharacterised BMCs are entering our views that await further investigations (Axen et al., 2014; Kerfeld et al., 2018). The expanding range of functions of BMCs suggested the adaptive nature of BMCs to cope with specific environmental niches (Chowdhury et al., 2014).

1.2 Carboxysome as a model for understanding BMC

To elucidate the BMC adaption in a native organism, we started our research on carboxysomes that have relatively straightforward, yet irreplaceable biological function. Unlike Pdu and Eut which are only required when supplied with particular nutrients, carboxysomes are fundamentally essential for carboxysome containing species to survive in natural environment. Removing or impairing carboxysomes would render cells incapable of CO_2 fixation in ambient air that leads to cell death (Price and Badger, 1989). Carboxysomes contain only two type of enzymes, the Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) (Shively et al., 1973) and the carbonic anhydrase (CA) (McGurn et al., 2016) that catalyse a linearised enzymatic conversion. While in Pdu and Eut, a broader set of enzymes, formed by PduC/D/E/P/Q/L and EutE/D/G, are sustaining a more complicated multi-step catalytic processes for their host organisms (Bobik et al., 2015). Based on the simple composition of enzymes and catalytic processes, the carboxysome is an ideal model to uncover the function and structural adaption within the host organism.

1.2.1 CO_2 fixation by carboxysomes

Rubiscos have both carboxylase and oxidase activity depending on the substrates provided (**Figure 1-2**). In carboxysomes, enzymatic properties of Rubisco are optimised to favour the carboxylase

activities. The CO_2 molecules are initially stored in the form of bicarbonate (HCO_3^-) which is actively transported and concentrated in the cytoplasm by inorganic carbon (Ci) transporter system (Price et al., 2008). The carboxysomal shells allow the entry of enriched HCO_3^- into the lumen. Inside the carboxysome, HCO_3^- is converted to CO_2 by CA, which can be fixed by Rubisco. Simultaneously, the other substrate Ribulose-bisphosphate (RuBP), a key component in the Calvin-Benson-Bassham (CBB) cycle, can enter the carboxysome lumen through the pores of the shell proteins therefore enable the catalytic reaction by Rubisco (Rae et al., 2013). The diagram of CO_2 fixation by carboxysomes is shown in **Figure 1-3**. The shell is suggested to form a CO_2 barrier that stops the diffusion or leakage of CO_2 from carboxysome lumens to the cytoplasm (Dou et al., 2008; Frey et al., 2016; Rae et al., 2013). Hence, the CO_2 molecules are retained in carboxysomes at high local concentrations, therefore enhance the carboxylation efficiency of Rubiscos. Without the enrichment of CO_2 in the lumen, O_2 can compete with CO_2 to bind Rubisco reaction centre, significantly reducing the efficiency of Rubisco carboxylation efficiency (Griffiths, 2006).

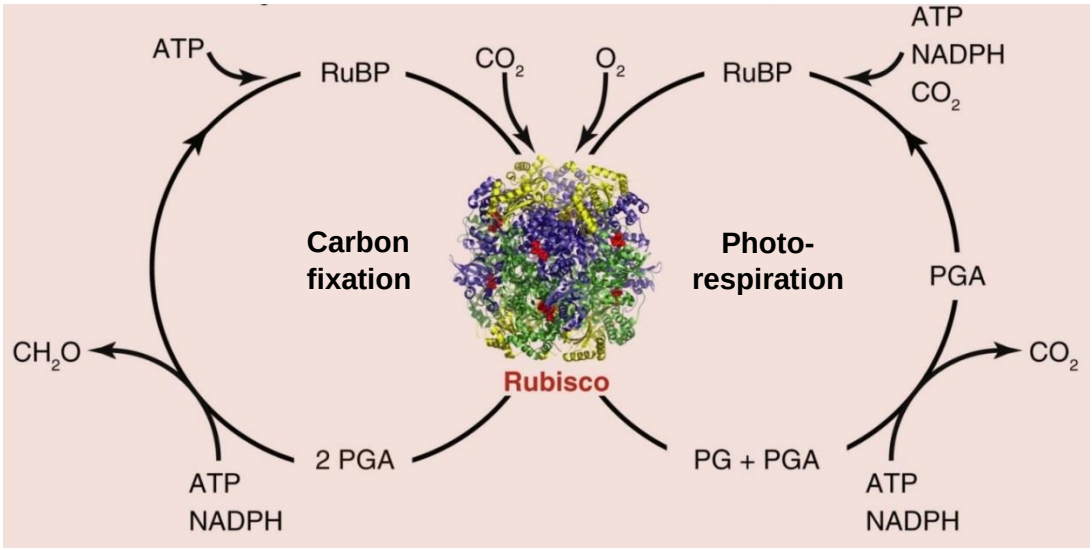


Figure 1-2. Carbon fixation and Photorespiration catalysed by Rubiscos (Kellogg, 2013). Hexadecameric Rubiscos could function as carboxylase and oxygenase depending on the substrate provided. CO_2 and O_2 are competing substrates to Rubiscos. CO_2 leads to carbon fixation in the CBB cycle while O_2 leads to the wasteful photorespiratory cycle. CH_2O represent fixed organic sugar in different forms. PGA short for 3-phosphoglycerate, PG short for 2-phosphoglycolate.

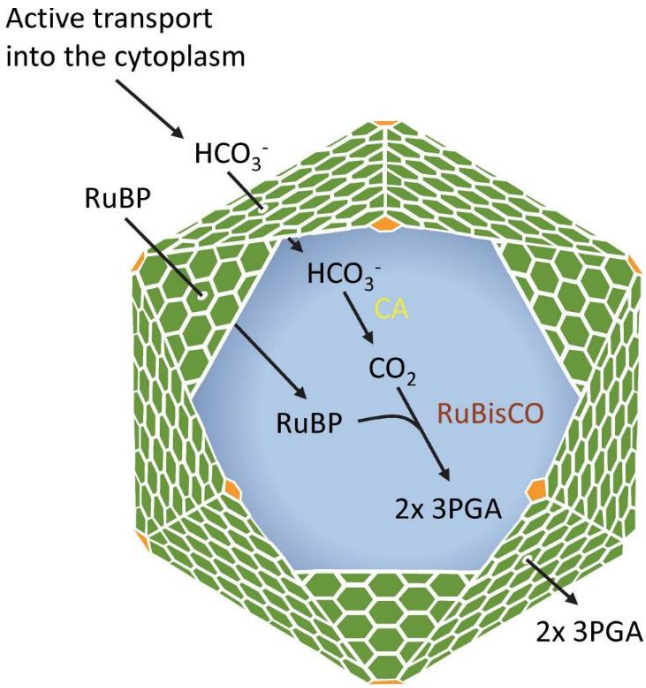


Figure 1-3. Diagrams of carboxysome functions (Bobik et al., 2015). Carboxysomes are capable of converting CO_2 and RuBP into 3PGA, which is the critical stage in the Calvin-Benson-Bassham Cycle (CBB cycle).

In higher plant, Rubiscos can utilise ambient levels of CO_2 in chloroplast in free forms without being sequestered compared with cyanobacterial Rubiscos (**Figure 1-4**) (Whitney et al., 2011). As a trade-off, plant Rubiscos are slower in CO_2 fixation, at a rate between 2 to 6 CO_2 molecules per Rubisco per second, while cyanobacterial Rubiscos could achieve a higher rate at 12 to 14. Even with higher specificity of CO_2 over O_2, most plants, typically C3 plants, still lose about 30 % of carbon fixation capacities to the undesired photorespiration in natural conditions globally (Zhu et al., 2010).

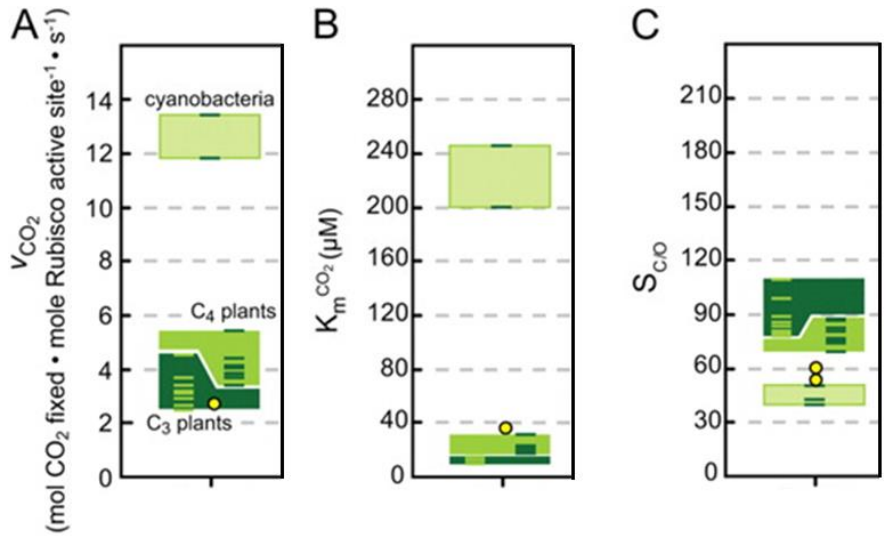


Figure 1-4. Comparative catalytic features of cyanobacterial and plant Rubiscos. The fixation rate as V_{CO2} (A), CO_2 affinity as K_m^{CO2} (B) and specificity of CO_2 over O_2 as $S_{C/O}$ calculated from $(V_{CO2}/K_m^{O2})/(V_{O2}/K_m^{CO2})$ (C). Circles coloured in yellow correspond to values for green algae. Figure adapted from (Whitney et al., 2011). (A) indicate faster turn-over of substrate for cyanobacteria compared to plants, while (B) suggest lower CO_2 affinity to enzyme and lesser competitive of CO_2 over O_2 for reaction.

The faster CO_2 fixation of cyanobacterial Rubiscos is contributed by multiple attributions supported through compartmentation. Free forms of cyanobacterial Rubiscos when not encapsulated by the shell are incapable of sustaining carboxylation activities in ambient levels of CO_2 (Galmes et al., 2014; McNevin et al., 2006). The high affinity for O_2 against CO_2 would eventually favour the wasteful photorespiration reaction through oxidation. The encapsulation of Rubisco and CA in carboxysomes elevate CO_2 concentrations, therefore, compensating the low CO_2 affinity. The carboxysome shells are also suggested to exclude O_2 hence prevent photorespiration and reduce the leakage of CO_2 out of the carboxysomes (Dou et al., 2008; Frey et al., 2016; Rae et al., 2013). Consequently, the local levels of CO_2, as well as the ratio of CO_2 against O_2 could be effectively elevated, further contributing to a favourable carboxylation conversion by Rubiscos.

1.2.2 Two types of carboxysomes

According to the types of Rubisco (form 1A and form 1B) encapsulated, two types of carboxysome, as α-carboxysomes and β-carboxysomes, have been categorised (Jakobson et al., 2017; Sommer et al., 2017). The independently evolved α- and β-carboxysomes are found mostly in oceanic and freshwater strains respectively (Rae et al., 2013). Although separately evolved, both types of carboxysomes share high degrees of similarities in protein composition and overall structure assemblies.

The protein components of α- and β–carboxysomes are listed in **Table 1-1**, where compositions for both types of carboxysome in model organism *Halothiobacillus neapolitanus* C2, a bacteria that fix CO_2 by their energy derived through oxidation of reduced sulphur compounds, and Syn7942 were compared. Homolog proteins achieve same core functions across two types of carboxysomes. On the surface of carboxysomes, CcmK2-K4 (Kerfeld et al., 2005; Samborska and Kimber, 2012; Tanaka et al., 2008; Tanaka et al., 2009) and CsoS1ABC (Tsai et al., 2007) that all belong to the BMC-H family cover the majority of shell facets. CcmO/P (Cai et al., 2013; Larsson et al., 2017; Marco et al., 1994) and CsoS1D (Klein et al., 2009) that belong to the BMC-T family are minor components assemble into the shell facets. Meanwhile, CcmP is believed to forms the gateway system of the shell, therefore, modulate the shell permeability (Kerfeld and Melnicki, 2016). CcmL (Tanaka et al., 2008) and CsoS4A/B (Cai et al., 2009) from the BMC-P family assembly 5-fold symmetric pentamers are deduced to cap the vertices of the polyhedron shells.

In carboxysome lumens, Rubisco and CA are the core proteins occupying the most spaces. For monomeric Rubisco, both form IA and IB types shared hexadecameric (L_8S_8) formation (Schneider et al., 1992). It is composed of eight large subunits (LSUs) coded by *cbbL* or *rbcL* and eight small subunits (SSUs) coded by *cbbS* or *rbcS*. The reaction core is formed through tetramerisation of four LSU dimmers and locates within the formed LSU L_8 assembly. Eight SSU complement the L_8 and form a L_8S_8 complex, further improve the overall catalytic efficiency of the reaction core (Spreitzer, 2003).

Table 1-1. Protein information of typical α- and β-carboxysomes in model organisms *Halothiobacillus neapolitanus* C2 and Syn 7942.

α-carboxysome		β-carboxysome		Coding gene localisation on genome
Protein	Description	Protein	Description	
CbbL	Rubisco large subunit	RbcL	Rubisco large subunit	Main operon
CbbS	Rubisco small subunit	RbcS	Rubisco small subunit	
CsoS1A	Hexamers, major shell facet	CcmK2	Hexamers, major shell protein	
CsoS1B				
CsoS1C				
CsoS4A	Pentamers, putative vertices of shell	CcmL	Pentamers, putative vertices of shell	
CsoS4B				
CsoS2	Structural protein	CcmM	Structural protein	
		CcmN		
CsoS3	Carbonic anhydrase	CcmO	Pseudo- hexamers; Putative edge shell proteins	
		CcaA	Carbonic anhydrase	Satellite loci
CsoS1D	Pseudo- hexamers Minor shell protein	CcmP	Pseudo- hexamers; Putative minor shell proteins	
CbbQ/CbbO	Potential Rubisco activase	CcmK3	Hexamers; Minor shell proteins	
		CcmK4		
		RbcX	Rubisco assembly chaperon	

The CAs coded by *csoS3* and *ccaA* (Long et al., 2007) in α- and β-carboxysomes come from ε-Class and β-class CA respectively (McGurn et al., 2016; So et al., 2004). These CAs have no significant sequential similarity but share similar catalytic properties in bicarbonate/CO_2 conversions (So et al., 2004). Such convergent evolution of CAs further supports the independent evolution of the carboxysomes.

Typical carboxysomes in different organisms share common features, but also retain several pronounced differences (Rae et al., 2013). Although shells and core enzymes proteins share high structural homology, the genetic contents are distinctive. The genetic composition of model α- and β-carboxysomes in organisms *Halothiobacillus neapolitanus C2* and Syn7942 are displayed in **Figure 1-5**.

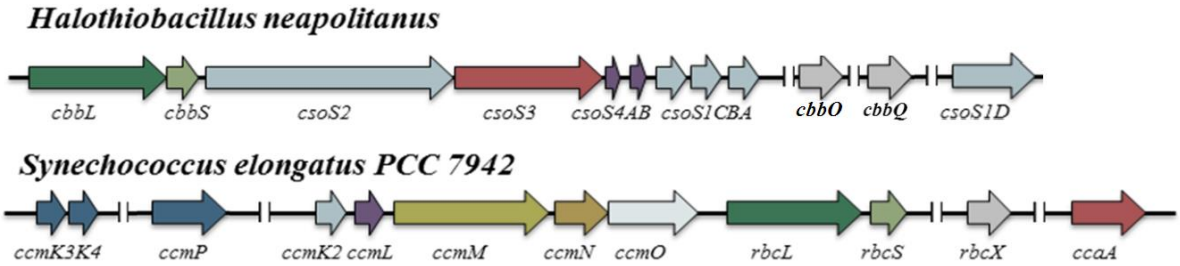


Figure 1-5. Distribution of carboxysomal genes for α- and β-carboxysomes on the genomic region from *H. neapolitanus* and *Synechococcus elongatus* PCC7942. Double-slash lines indicate gaps over distant genomic regions separated by other genetic contents.

For α-carboxysomes, all genes required for carboxysome assembly and function (9 out of 11 genes) are located over a single *cso* operon (Bonacci et al., 2012). The additional satellite carboxysomal genes *csoS1D* (Klein et al., 2009) and *cbbQ/cbbO* (Sutter et al., 2015; Tsai et al., 2015) are capable of further enhancing the carbon fixation capabilities. The genetic content of β-carboxysomes has higher complexity than that of α-carboxysomes. Besides the major *ccm* and *rbc* operons that encode carboxysomal genes from *ccmK2* to *rbcS*, four distant loci have been identified encoding BMC-T and BMC-H shell proteins, as well as the crucial CA enzyme (CcaA) and Rubisco chaperone RbcX (**Figure 1-5**).

Unlike α-carboxysomes which functions can be primarily retained (less than 50% decrease in CO_2 fixation capacities) without proteins coded by satellite genes such as *csoS1D* and *cbbQ* (Bonacci et al., 2012), function integrity of β-carboxysomes that lack satellite proteins is severely impaired: carboxysomes lacking both CcmK3 and CcmK4 are unable to fix CO_2 in ambient air, rendering the mutant strain to have a High-CO_2 Requirement (HCR) phenotype (Rae et al., 2012). The greater involvement of multiple operons in β–carboxysomes indicates higher degrees of structural and functional complexities. Meanwhile, it also suggests distinctive modulation of β-carboxysomes in freshwater habitats (Whitehead et al., 2014). However, current understandings of such potential adaption are mainly predictions inferred from indirect evidence (i.e. transcriptomic profile). Direct evidence over the *in vivo* carboxysome modulation adapting to environmental changes remains to be investigated.

Moreover, the aggregating strategy for Rubisco enzymes which form the interior matrix is drastically different in the two types of carboxysomes. In α-carboxysomes, a loose and irregular L_8S_8 hexadecameric complexes aggregation (shown in **Figure 1-6A** and **C),** driven by structural protein CsoS2 through putative binding affinities are widely acknowledged (Cai et al., 2015a). In β-carboxysomes, the interior assembly is mediated by structural proteins CcmM and CcmN, shown in **Figure 1-6B**. CcmN serves as a linker through the encapsulation peptide on C-terminus (Aussignargues et al., 2015) that associates CcmK shell proteins and an N-terminal domain that associates with CcmM (Kinney et al., 2012). CcmM exists in two forms (short isoform CcmM35 and long isoform CcmM58) by alternative translation from start codons locate at the beginning and middle of the gene coding sequence (Long et al., 2007). Both the long and short isoforms contain three RbcS-like homolog domains which are believed to replace RbcS in the Rubisco hexadecameric L_8S_8 assemblies, leading to the formation of a more regular array-like compact-packing paracrystalline resemble $L_8S_5M_3$ stoichiometry (Long et al., 2011). While the long isoform CcmM58 processes an additional N-terminal domain that binds CcmN, therefore, enables the encapsulation of shells (Long et al., 2007; Long et al., 2011; Long et al., 2010).

The natures of loose/irregular and compact/regular packing of core enzymes (**Figure 1-6C**) eventually lead to different assembly strategies between α- and β-carboxysomes. Recent studies documented that *de novo* assembly of β-carboxysomes exploits the core-based "inside out" model: Rubisco and CcmM forming the core first, followed by the encapsulation of shell proteins, mediated by CcmN and CcmK (Cameron et al., 2013; Chen et al., 2013). Such compact packing model could support more Rubiscos in each carboxysome, resulting in higher Rubisco packing density (Rae et al., 2013) and greater carboxysome diameters ranging from 200nm to 400nm (Cai et al., 2015a). Moreover, the dense core of β-carboxysomes is suggested to serve as the backbone of the overall structures that contribute to the integrity of carboxysome structures (Kerfeld and Melnicki, 2016; Rae et al., 2013). While in α-carboxysomes, the overall structural integrity is not supported by the loose interior, but mainly through shell protein interactions (Kerfeld and Melnicki, 2016). Moreover, an overall smaller

diameter of α-carboxysomes, ranging from 100nm to 160nm (Heinhorst et al., 2014), further limited the enzyme loads per carboxysome compared with β-carboxysomes.

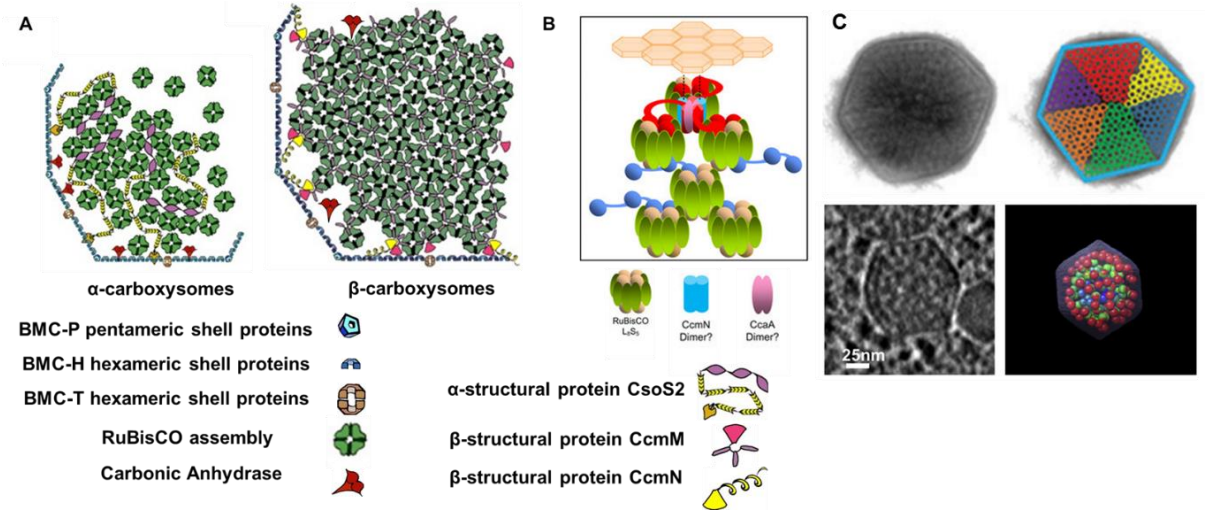


Figure 1-6. Current understanding of the structural basis for α- and β-carboxysomes, adapted from (Dai et al., 2018; Faulkner et al., 2017; Kerfeld and Melnicki, 2016; Rae et al., 2013). A. diagrams of α- and β-carboxysome assemblies, legends shown below. B. Diagrams of β-carboxysomes assemblies with $RbcL_8RbcS_5CcmM_3$ formation and putative CcmN and CcaA location. C. EM images and fitted Rubiscos of α- and β-carboxysomes (bottom and top respectively) indicate differences of packing patterns in the lumens. The circles in different color indicate clusters of Rubiscos with less ordered/ordered formation for α- and β-carboxysomes.

Last but not least, Rubisco chaperones RbcX (Emlyn-Jones et al., 2006; Occhialini et al., 2016) or Rubisco activase (Rca) CbbO/CbbQ (Sutter et al., 2015; Tsai et al., 2015) are only found in partial genomes of α- and β-carboxysomes containing organisms (Zarzycki et al., 2013). However, these chaperones and activases are not commonly shared by carboxysomes. Particularly for the two model strains showed in **Table 1-1**. The chaperone *cbbX* gene corresponding to *rbcX* is not found in *Halothiobacillus*. Instead, a pair of genes called *cbbO*/*cbbQ* has been identified, which product are capable of enhancing Rubisco enzymatic performance through the formation of hetero-oligomers when co-expressed with Rubisco heterogeneously (Tsai et al., 2015). Meanwhile, Rca corresponding to CbbO/CbbQ is not present in Syn7942 genome. The existence or loss of Rubisco helper proteins complement diverse Rubiscos in different organisms to sustain desired level of CO_2 fixation

properties and are not considered to affect structural integrity of carboxysomes (Emlyn-Jones et al., 2006; Tanaka et al., 2007; Zarzycki et al., 2013), though further investigation on their roles are needed.

Overall, α- and β-carboxysomes are both capable of overcoming the disadvantages of the cyanobacterial Rubiscos by intelligently designed assemblies, achieving faster fixation rate compared with their homologs in higher plants.

1.2.3 Synthetic engineering of carboxysomes

In plants, Rubiscos are exposed in chloroplasts stroma where O_2 are generated continuously from the photosynthetic complexes locate on the thylakoid membranes in close proximity (Harris and Königer, 1997). Particular plants have evolved different carbon concentrating strategies: C4 plants evolve specific bundle-sheath cells, which capable of housing the separated step of carboxylation from substrate pumped by initial carbon fixation catalysed by PEPCase (Furbank et al., 1989; Gowik and Westhoff, 2011). Such mechanism is also considered as a type of carbon concentrating mechanism (CCM); while plants grown in arid climate utilise Crassulacean Acid Metabolism (CAM) that separate major CO_2 fixation activity into the night from photosynthesis activating during day times (Keeley et al., 2003). However, the majority of crop plants, especially the most cultivated crops such as wheat and rice possess C3 pathway without any form of CCM (Ducat and Silver, 2012).

The advantageous CO_2 fixation properties (**Figure 1-4**) through Rubisco encapsulation in carboxysomes provides an alternative strategy for improving CO_2 fixation in higher plants. The hypothesis that significant improvement could be achieved by introducing carboxysomes into chloroplast has attracted tremendous interest in engineering the CO_2-fixing organelle in higher plants. In this field, α-carboxysomes seems more suitable to be manipulated compared with β-carboxysomes that have higher degrees of complexity. For instance, partial carboxysome structures have been reproduced heterologously in forms of protein sheets, nanotubes or entire coreless polyhedron shells

(Fan and Bobik, 2011; Pang et al., 2014; Parsons et al., 2010; Sutter et al., 2017). Furthermore, without much modification to promoters and coding sequences, fully functional α-carboxysomes have been successfully generated in *E. coli* and a gram-positive bacterium *Corynebacterium glutamicum* (Baumgart et al., 2017; Bonacci et al., 2012).

Even though α-carboxysomes seems to be a better choice for engineering based on the above findings, β-carboxysome synthesis still attracts great attention to researchers due to the close nature of Rubisco with higher plants as well as the greater encapsulation capacity compared with α-carboxysomes (described in **section 1.2.1**). As a proof of concept, researchers have genetically engineered partial components of β-carboxysomes (Rubisco and shell respectively) into tobacco chloroplast (Hanson et al., 2016; Lin et al., 2014a; Lin et al., 2014b). These pioneer works have shown the successful assembly of β-cyanobacterial Rubiscos and shells in plant chloroplasts, giving us more faith in reaching the final goal of introducing complete β-carboxysomes for enhanced carbon fixation in higher plants.

More challenges were found when trying to reproduce entire β-carboxysomes heterologously (Cai et al., 2016; Gonzalez-Esquer et al., 2015; Hanson et al., 2016; Lin et al., 2014a). The fundamental difficulty is to design vectors to express carboxysomal genes over the multi operon at correct ratios. Up to date, it remains unknown what the exact expression profile is for carboxysomal genes under the native regulation, and a step further, the final stoichiometry of carboxysome assembly. Besides, it is suggested that β-carboxysomes assembly factor Raf1 might also be required to promote initial Rubisco assembly in *E. coli* cells heterologously (Kolesinski et al., 2014). Till now, no full β-carboxysome structure has been successfully reproduced outside native host cells, where the latest attempt that was published recently showed carboxysome-like structures probably due to uncontrolled protein stoichiometry and organisation (Fang et al., 2018).

An alternative strategy that simplifies the complicated β-carboxysomes to a reduced model has also been tested: carboxysomes with an artificial chimeric hybrid protein, named as CcmC, that replace

native CcmM, CcaA, and CcmN were successfully generated in native cyanobacteria host (Gonzalez-Esquer et al., 2015). The chimeric carboxysome can replace native carboxysome to sustain CO_2 fixation in ambient air. However, reduced organelle diameter, as well as uneven Rubisco loads in carboxysomal lumens, are also observed. The reduced sizes and packing density of Rubiscos in this chimeric carboxysome suggest the importance of maintaining proper stoichiometry of structural protein CcmM, CcmN and CcaA to obtain optimised carboxysome assemblies when expressed heterologously.

Overall, in the author's view, the key to overcoming the difficulties in β-carboxysome engineering requires a better understanding of carboxysome stoichiometry, assembly as well as their regulations within the native host.

1.3 The model organism Syn7942 for β-carboxysome study

1.3.1 Advantages of Syn7942 as a model organism

Cyanobacteria, once known as blue-green algae, are primitive aquatic prokaryotes among the oldest organisms on our planet. They are widely adapted species and fundamental members in almost every current ecosystem (Dvornyk et al., 2003). These ancient prokaryotes are the originators of photosynthesis, responsible for 25% of the global carbon fixation in the oceanic environment (Bullerjahn and Post, 2014; Flombaum et al., 2013). *Synechococcus elongatus* PCC7942, previously known as *Anacystis nidulans* R2, has been extensively selected as a model strain to study carbon assimilation (Tchernov et al., 2001), acclimation to environmental changes (Bustos and Golden, 1992; Schwarz and Grossman, 1998; Tsinoremas et al., 1994) as well as bacterial circadian clocks (Cohen and Golden, 2015; Golden, 2003; Swan et al., 2018). This rod-shaped organism is superior for microscopic studies over carboxysomes due to its distinctive carboxysome contents (three to four carboxysomes per cell) in clear and ordered cytosolic spaces and well-separated subcellular localisations (Savage et al., 2010; Yokoo et al., 2015). On the other hand, Syn7942 is reliably transformable by exogenous DNA naturally (Shestakov and Khyen, 1970), providing us with easy

access for carboxysome genes mutagenesis. Together with the most comprehensive and well-documented gene and protein information in the database, Syn7942 serves as a perfect model organism for native carboxysome investigation *in vivo*.

1.3.2 Traits of Syn7942 carboxysomes and β-carboxysomes in other organisms

Through evolution, The protein compositions of β-carboxysomes in different organisms have adapted to different habitats diversely (Badger et al., 2006). It is essential to consider how Syn7942 carboxysomes differ from other β-carboxysomes: What are features commonly shared among all β-carboxysomes? What features are Syn7942-specific? Overall, carboxysomes in Syn7942 contain all the representative proteins in each category, including CcmK2 to RbcS coded by genes located within the main *ccm* and *rbc* operons, together with the CcaA, CcmK3/K4 and RbcX encoded by genes located on the satellite loci.

The most considerable difference against other β-carboxysomes is that an additional major shell protein CcmK1 was absent in Syn7942. CcmK1 and CcmK2 are highly conserved structurally, with CcmK1 containing a ten amino acids long C-terminal extenstion compared with CcmK2 (Sommer et al., 2017). However, truncation of the C-terminus in both CcmK1 and CcmK2 does not prevent regular shell assembly, suggesting non-essential roles of extended C-terminus on CcmK1 (Cai et al., 2016). Instead, CcmK1 is theoretically considered to be a redundant duplicate of CcmK2 that plays a role in balancing the overall shell stoichiometry at translational level (Sommer et al., 2017). The validation of CcmK1/CcmK2 stoichiometry and their quantitative relevance is beyond the grasp of this thesis.

Another difference with most β-carboxysomes is RbcX. The *rbcX* that typically found between *rbcL* and *rbcS* in the *rbc* operon (Onizuka et al., 2004; Tanaka et al., 2007) is alternatively located at a distant satellite loci in Syn7942. The chaperonic role which was essential in β-carboxysomes (Saschenbrecker et al., 2007) in other organisms are not required for carboxysome functions (Emlyn-

Jones et al., 2006). The contradictive findings in RbcX functions suggested some specific functions of RbcX in Syn7942, which require further characterisation.

Lastly, the CA coded by *ccaA* was missing in several β-carboxysomes containing species. Instead, the catalytic function was fulfilled by the N-terminal domain of CcmM which also process γ-CA activity (Pena et al., 2010). The deletion of CcaA allowed cells to exhibit an HCR phenotype, indicating that the catalytic function was still primarily sustained by the CcaA in Syn7942 (So et al., 2002).

1.4 Subcellular positioning of carboxysomes

The intracellular organisation is of great importance for smooth metabolic and biosynthetic activities in bacteria (Rudner and Losick, 2010; Shapiro et al., 2009). Association of carboxysomes with cytoskeletal networks has been elucidated (Bobik, 2006; Yeates et al., 2011). Mainly, Syn7942 has evolved a comprehensive subcellular positioning system to organise the crowded and dynamic viscera of the cells. Crucial subcellular components such as carboxysomes and chromosomes at low copy numbers are spatially distributed within cells, ensuring the equal heritance to daughter cells (Jain et al., 2012; Savage et al., 2010). A ParA-like filament based model has been proposed, indicating that carboxysome positioning is controlled by the cytoskeletal system (Savage et al., 2010). Experimental evidence has been reported supporting the above theory. Cytoskeletal proteins such as ParA and MreB have been identified in the fraction of isolated carboxysomes (Faulkner et al., 2017).

Meanwhile, a previous study by electron microscopy shows the peripheral arrangement of carboxysomes near the thylakoid membranes (McKay et al., 1993). The author also suggested functional advantages in Ci utilisation when carboxysomes are closer to the Ci uptake system over membranes. The discrepancies found in carboxysome localisation require further investigation.

Nevertheless, carboxysome mobility is an energy consuming process. The observed localisation dynamics is not limited to the time point during cell divisions. Therefore, such energy consuming

repositioning should have functions besides separation of carboxysomes to daughter cells during cell divisions (Jain et al., 2012). From **Chapter 3** to **Chapter 5**, the localisations of carboxysomes were investigated under different environmental conditions, with the aims of finding clues for the suggested functions. Even though such subcellular positioning mechanism might be mostly unique to Syn7942, the organelle positioning mechanism in elongated cells might serve as an excellent reference to elucidate how cells manage organelles within the limited subcellular space.

1.5 Ci uptake systems

In cyanobacterial cells, the Carbon Concentrating Mechanisms (CCMs) is consist of carboxysomes and Ci uptake system that work together to sustain the carbon fixation activity (Badger and Price, 2003). The Ci uptake system locates on plasma membrane is capable of transport the substrate HCO_3^- passively into the cytoplasm (Price et al., 2008). In general, multiple types of Ci transporter have been identified, including an ATP-dependent high-affinity transporter BCT1 (Omata et al., 1999), electrochemical Na^+ coupled medium affinity transporters SbtA and BicA (Price et al., 2004; Shibata et al., 2002), and two isoforms of reduced β-nicotinamide adenine dinucleotide phosphate (NADPH) or Ferredoxin (Fd) dependent NADPH NDH-I complexes (Maeda et al., 2002; Shibata et al., 2001). BCT1, SbtA and BicA are capable of transport extracellular HCO_3^- into cytoplasm while NDH-I converts cytosolic CO_2 into HCO_3^-, further reduced the concentration of cytosolic CO_2 which promote the inward diffusion of CO_2 across the plasma membrane (Burnap et al., 2015), as shown in the upper panel of **Figure 1-7**. The Ci uptake system functionally enables the concentration of Ci pools in the cytoplasm to reach a strikingly thousand-fold higher than that of the environmental Ci level (Price et al., 1998; Woodger et al., 2005), therefore elevates CO_2 fixation efficiency and reduces unwanted photorespiration for carboxysomes (Eisenhut et al., 2008; Hagemann et al., 2013; Schwarz et al., 1995).

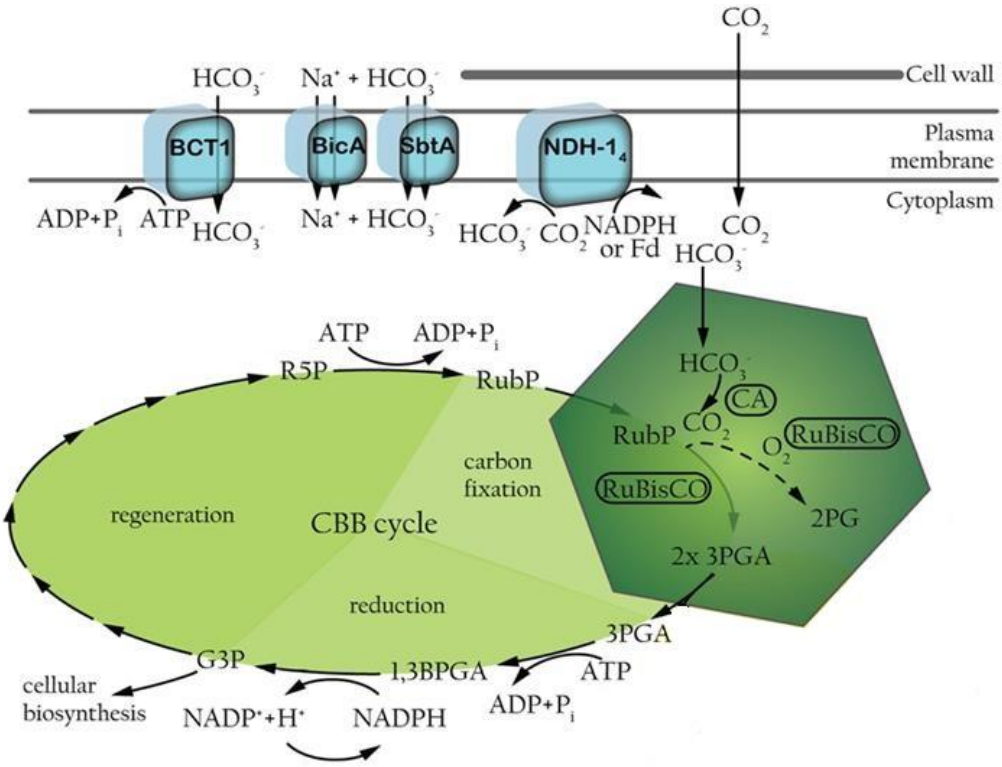


Figure 1-7. Modulation of carboxysomes interrelated with Ci transport system and CBB cycle. The Ci transporter, CBB cycle are interlinked in cyanobacterial cells. Adapted from (Montgomery et al., 2016). The crucial step of carbon fixation is contained within carboxysomes.

1.6 Carbon fixation and photosynthetic electron flow

The dominant oxygenic phototroph by cyanobacteria laid the foundation of atmosphere composition 3.5 billion years ago (Schopf and Packer, 1987). Moreover, cyanobacterial photosynthesis is sustaining the CO_2 fixation and O_2 regeneration in the atmosphere, contributing to the sustainable ecological environment nowadays. In cyanobacteria, solar energy from the sunlight is effectively converted by photosynthetic machinery locates on thylakoid membranes (Liu, 2016). The harvested energy is then utilised to support metabolic activities fundamentally required for cell survival and reproduction (Asayama, 2006; Ohbayashi et al., 2013). The utilisation of solar energy is achieved through the generation and transport of electrons which provide essential driving force for the synthesis of energy carriers such as ATP and NADPH.

In the model organism Syn7942, extensive research over photosynthesis is available, providing us with the well-documented background to investigate carboxysomes within the cellular environment. The photosynthetic machinery in Syn7942 mainly consists of the light-harvesting phycobilisome, Photosystem I (PSI), Photosystem II (PSII), cytochrome (Cyt) b6f and ATP synthase (ATPase), illustrated in **Figure 1-8**. The solar energy is first absorbed by phycobilisomes through light harvesting processes. PSII catalyses the oxidation of H_2O molecules, splitting them into H^+ and O_2, generating electrons which are donated to the Plastoquinone (PQ) pool. Electrons from the PQ pool are subsequently passed onto cytochrome (Cyt) b_6f complexes to pump protons across the thylakoid membrane and create an electrochemical gradient. When protons move down the gradient, ATPase obtains a spinning motion which bond ADP and Pi into ATP.. Over PSI, electrons are also generated and transferred away by intermediate electron carrier called ferredoxin (Fd). On the other hand, electrons are transferred further down the linear electron transport chain to PSI. Ferredoxin then carries the electrons to ferredoxin $NADP^+$ reductase (FNR) to reduce $NADP^+$ to NADPH, or returns the electrons to the PQ pool, forming the cyclic electron transfer chain.

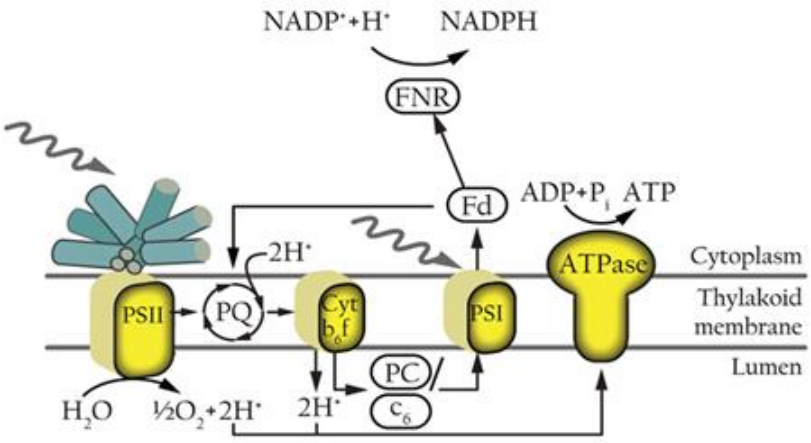


Figure 1-8. Components of cyanobacterial photosynthetic complexes and electron transporters. Adopted from (Montgomery et al., 2016).

The photosynthesis converts solar energy into usable forms of ATP and NADPH, providing the driving force to the Calvin-Benson-Bassham Cycle (CBB Cycle) that fix CO_2 by carboxysomes (Drews and Niklowitz, 1956). In general, the CBB cycle (Shown in **Figure 1-7**) could be divided into

three phases: the carboxylation, reduction and regeneration. In cyanobacteria, the initial carboxylation phase in the carboxysome is catalysed by Rubisco where the RuBP combines with CO_2 at 1:1 ratio. The generated 3PGA is then catalysed into glyceraldehyde 3-phosphate (G3P) with the energy provided by adenosine triphosphate (ATP) and reduced β-nicotinamide adenine dinucleotide phosphate (NADPH) regenerates through ferredoxin-$NADP^+$ reductase (FNR) from oxidised β-nicotinamide adenine dinucleotide phosphate ($NADP^+$) and proton (shown as H^+ in Figure 1-7). The G3P could be utilised by cellular biosynthesis providing essential building blocks for of biomass accumulation, at the same time, regenerated into ribose-5-phosphate (R5P), converted to ribulose-5-Phosphate (Ru5P) that is then coverted by ATP back into RuBP therefore regenerated through the CBB cycle.

Overall, ATP and NADPH are generated through the electron flow in photosynthetic light-dependent reactions. For cyanobacteria, light availability, particularly the intensity, is fluctuated on the daily and seasonal basis (Allahverdiyeva et al., 2015). Indeed, the light intensity is considered as the rate-limiting factor for PSII, PSI and cytochrome (Cyt) b_6f (Vermaas, 2001). Therefore, the light intensities ultimately determined the rates of ATP and NADPH synthesis.

1.7 Methods used in carboxysome studies

1.7.1 Transcriptomic approaches

The development of high-throughput RNA-sequencing enables us to have a better understanding of carboxysome biogenesis within cellular metabolism. Transcriptomic data assist the recognition of carboxysomal genes through clustering of co-expression profile, particularly useful when the gene contents are located distant over the genomes. For instance, in carboxysome studies, the distant *csoS1D* was found to co-express with known carboxysomal genes in *cso* operon (Klein et al., 2009).

Meanwhile, transcriptomic data also provide complementary evidence suggesting variation of carboxysomal contents. However, transcriptomic data alone is insufficient to support content changes

on protein level with the involvement of post-transcriptional regulations. As we lack the evidence of variation on protein level, the post-transcriptional regulation in carboxysomes remains unclear. In this thesis, we documented the carboxysomal content variation on functional protein levels, which would be crucial to uncover the regulations of carboxysomal stoichiometry and structural assembly essential for future carboxysome engineering.

1.7.2 Proteomic approaches

The application of proteomics in the past two decades has impacted BMC studies significantly, providing essential methods for carboxysomal protein validations as well as quantifications. In general, researchers have been adopting two methods widely, immunoassay and mass spectrometry (MS). Classic immunoassay using specially raised antibody is widely considered as direct evidence validating the presence of the target protein. For β-carboxysomes, antibodies have successfully raised against carboxysomal proteins in separate works: CcmK (which recognise all variants of CcmK homologs including CcmK2/3/4) (Rae et al., 2012), CcmO (Cai et al., 2016), CcmM (recognise both isoforms CcmM35/58) (Long et al., 2007), CcaA (Long et al., 2007), RbcL (Long et al., 2007), RbcS (Rae et al., 2012), CcmL (Cai et al., 2016) and RbcX (Emlyn-Jones et al., 2006), while antibodies specific for CcmN, CcmP are not yet available. Researchers have adopted immunoblot quantification to investigate stoichiometries from isolated carboxysomes or cell lysates (Long et al., 2007; Long et al., 2011).

However, the effectiveness of quantification is limited by sensitivities of antibodies. The sensitivities of anti-carboxysomal antibodies are relatively low compared with commercial antibodies raised against non-bacterial originated proteins such as GFPs/YFPs. Some minor carboxysomal proteins such as RbcX (Emlyn-Jones et al., 2006) and CcmL (Cai et al., 2016) are so less in abundances that cannot be detected in cell lysates. Enrichment of carboxysomes through purification and isolation can overcome the sensitivity issue, but for proteins such as CcmN, CcmP and RbcX, detection cannot be achieved even in isolated carboxysomes, raising the possibility of potential destruction of carboxysome structures during isolation that leads to the release of these proteins from enriched

fractions (Faulkner et al., 2017). An alternative strategy is to enhance the recognition of target BMC protein by fusing tags with better specificity and sensitivity of antigen-antibody. Researchers have chosen poly-histidine tag (His-tag) due to its small size and cleavage capability after purification for carboxysomal protein recognition and relative quantification (Cai et al., 2013; Rohnke et al., 2018; Tanaka et al., 2009).

Besides, relative quantification by high-throughput MS has also been adopted, giving researchers more power of detection beyond the limitation of antibodies (Long et al., 2005; Mayer et al., 2016; Zarzycki et al., 2017). The stoichiometry of RbcL and RbcS detected in isolated carboxysomes suggest a likely L_8S_5 assembly (Long et al., 2011); While CcmM58 and CcaA content are correlated linearly between ambient air and 2% CO_2 growth conditions, suggesting defined assembly symmetries between these two proteins. However, due to the multiple numbers of carboxysome in cells, the stoichiometry of individual carboxysome cannot be determined conclusively by both immunoblot, and MS approaches from cell lysates, while isolation of carboxysomes suffers from possible protein losses.

In this thesis, we extend the usage of immunoblot for quantification purposes by swapping native protein with FP-tagged versions in the original genomic regions. In this way, the expression and translation of FP-tagged proteins are under native control, closely resemble the conditions for their wild-type counterparts. Moreover, the universal FP-tags serve as a standard, would enable quantitatively comparison of different tagged proteins which were unable to achieve through special-raised antibodies, providing relative quantification that assists in stoichiometry studies shown in **Chapter 4** and **Chapter 5**.

1.7.3 Bioinformatic approaches

Comparative genomics trigger the expanding of BMC studies across different organisms since the first complete sequencing of bacterial genome (Fleischmann et al., 1995; Fraser et al., 1995). The first BMC being discovered, the carboxysome was found primarily due to the constant existance in cells

with its iconic polyhedron shapes in cells when observed under microscopy back in 1956 (Drews and Niklowitz, 1956). With the advance in genomic sequencing, homologs of carboxysome shell proteins were identified in propanediol utilising Salmonella (Chen et al., 1994), eventually leads to the discovery of new types BMCs, the Pdu microcompartment (Crowley et al., 2008). Unlike carboxysomes that are required in cells under general cultivation conditions, Pdu microcompartments are only induced in the host cells when supplying with specific nutrient propanediol.

Furthermore, the discovery of satellite carboxysomal genes such as *ccmP*, *csoS1D ccmK5/K6* (Cai et al., 2013; Klein et al., 2009; Sommer et al., 2017) and complicated genomic content for same carboxysomal genes across different species through comparative genomic studies (Sommer et al., 2017) leading to the newly developed concept called BMC superloci (Kerfeld et al., 2018). Beyond the field of carboxysomes, new class of BMC such as glycyl radical enzyme-associated microcompartment (GRM) (Zarzycki et al., 2015), the Planctomycetes and Verrucomicorbia microcompartment (PVM BMC) that capable of degradating the L-fucose and L-rhamnose (Erbilgin et al., 2014) have been identified through comparative genomics studies. To date, homologs of BMC shell proteins have been identified in 23 out of 29 bacteria phyla, suggesting the universal adoption of BMC encapsulation systems among different bacteria species (Axen et al., 2014).

In conclusion, comparative genomic approaches are powerful in expanding known information across different species for BMC identifications independent of experimental characterisations and will continue to be crucial in the discovery of novel BMCs with increasing numbers of sequenced bacterial genomic entries in future.

1.7.4 Visualisation of carboxysomes

1.7.4.1 X-ray crystallography

Thousands of proteins assemble into a functional BMC unit. Instead of trying to characterise the entire BMC structure, strategy aiming at solving the structure of individual building block was initiated a decade ago, starting with the most studied shell protein CcmK by X-ray crystallography

(Kerfeld et al., 2005). Through analysing the diffraction profile of X-rays hitting on the crystallised proteins, three-dimensional photographs of atoms and the length as well as the types of chemical bonds could be interpreted, which eventually lead to the reconstruction of a three-dimensional model of the protein at atomic levels, as shown in **Figure 1-1**.

Considerable efforts have been taken to solve the structure of each BMC protein in the past decade, providing us with the fundamental structure models. In summary, the structures for all carboxysome shell proteins, at least one for each type in BMC-T, BMC-H and BMC-P have been successfully solved (Cai et al., 2015a; Larsson et al., 2017; Sutter et al., 2017). The functional assemblies of hexamer, pentamer are therefore confirmed. Besides, crystal structures for partial internal proteins have also obtained. For instance, Rubisco (Bracher et al., 2011), CcaA (McGurn et al., 2016), CcmM (Pena et al., 2010) and RbcX (Tanaka et al., 2007) structures are available in protein database solved by X-ray crystallography. Models obtained for different homologs across species indicate strong structural homology even when sequence similarities are far lesser conserved (Sommer et al., 2017).

Moreover, these high-resolution structures could be further analysed by computational simulations to predict their functional relevance. For instance, the permeability of BMC shell proteins was illustrated with simulated small molecules such as HCO_3^-, CO_2 and O_2 which suggest preferable permeability to HCO_3^- compared with CO_2 and O_2 inferred from the calculated free energy required to pass through the central pore of hexameric protein (Mahinthichaichan et al., 2018; Park et al., 2017). Following a reductionist approach, the dynamic assembly processes of BMC are also reproduced by simulations using simplified mimics with properties such as interaction strengths between shell subunit-subunit, shell subunit-cargo protein and cargo-cargo protein suggested from reported X-ray structures (Perlmutter et al., 2016; Rotskoff and Geissler, 2018).

However, the crystallography of some BMC structural proteins remains problematic. Particularly in carboxysomes, the structures of CcmN and CcmO in β-carboxysomes, as well as CsoS2 in α-carboxysomes remain unsolved. Sequence analysis of these proteins suggests flexible 3D formations,

therefore, lead to unsuccessful protein crystallisation. Another question raised is that whether individually crystallised structure represents the native biological assembly. A double-layered CcmK2 formation is observed from crystallisation data (Samborska and Kimber, 2012). However, this double-layered formation cannot be confirmed through Electron microscopy (EM) and Atomic force microscopy (AFM), where single layers of shell assembly are deduced (Faulkner et al., 2017; Sutter et al., 2016).

1.7.4.2 Electron microscopy

Electron microscopy (EM) has provided the first evidence of BMC existing in bacteria cells (Drews and Niklowitz, 1956). Furthermore, with the establishment of BMC isolation protocol, structural details regarding the arrangement of interior proteins and shells were uncovered (Faulkner et al., 2017; Iancu et al., 2007). For carboxysomes, the heterogeneity in BMC sizes and different interior assembly could be observed through EM (Faulkner et al., 2017; Iancu et al., 2007). Likewise, structural changes caused by knock-out of BMC protein of interest *in vivo* could be detected, providing functional references for the target protein (Cai et al., 2009; Cai et al., 2013; Parsons et al., 2010; Rae et al., 2012).

In recent years, a new technique called Cryogenic electron microscopy (cryo-EM) which could generate high-resolution structure that rivals the X-ray crystallography directly through native assemblies (Agirrezabala et al., 2015; Jiang and Tang, 2017; Lee and Gui, 2016; Sirohi et al., 2016) has been introduced from virus structure studies to BMC studies. Moreover, improvement in cryo-EM techniques over the past 10 years has reveal more paritial structures regarding individual rubisco and its assembly in carboxysome lumens (Dai et al., 2018; Schmid et al., 2006). The combination of X-ray crystallography together with cryo-EM generate high-resolution 3D structures of BMC shells assembly, revealing the basic principles such as the orientation of concave/convex as well as the assembly patterns of BMC-H, BMC-P and BMC-T proteins in an artificially expressed 6.5-MDa BMC (Sutter et al., 2017).

However, the disadvantages of EM methods are also significant. In the study of β-carboxysomes, researchers cannot depict more details such as the arrangement of homolog proteins from the minor structural differences through traditional EM. Meanwhile, the principle of adopting high-resolution imaging by cryo-EM requires BMCs to be enriched at high purity level, which require further optimising to exclude filamentous proteins such as ParA/MreB in the isolated fractions (Faulkner et al., 2017). Also, as an imaging technique based on fixed sample, EM cannot provide any aid to investigate the *in vivo* dynamics of structural changes that have been suggested to be a crucial feature of BMCs shells (Faulkner et al., 2017; Sutter et al., 2016).

1.7.4.3 AFM microscopy

Atomic force microscopy (AFM) enables the profiling of sample surface at atomic level by contacting by a probe which can magnify the profile of surface at scales of nanometres. AFM was initially designed to investigate flattened samples such as material surfaces in physics (Zhong and Yan, 2016) or biological membrane surfaces such as thylakoid membranes (Chuartzman et al., 2008; Liu and Scheuring, 2013). New application of AFM focusing on non-flattened material, particularly the virus (Kuznetsov and McPherson, 2011; Marchetti et al., 2016) stimulated the application of AFM on BMC studies (Rodriguez-Ramos et al., 2018). Up to date, several studies on sheet-like BMC shell proteins as well as intact BMCs have been done, providing insight into the uncharacterised properties of BMCs: the BMC shell hexamers can relocate over the assembled sheets dynamically, suggesting flexible protein-protein interactions (Garcia-Alles et al., 2017; Sutter et al., 2016), while intact BMCs are also flexible and have soft mechanical properties under indentation (Faulkner et al., 2017).

The most significant advantage of AFM is that the BMC structure can be visualised at the atomic level in a near-native condition. With a high-speed imaging variant of AFM, fast formation changes could be detected in a matter of seconds, capturing the dynamics effectively. However, the AFM also depends strongly on the isolation and purification of BMC samples. Contaminants in the imaging samples would increase background noises, leading to reduced imaging resolution. Meanwhile, the

artificial *in vitro* environment that mimic the *in vivo* status such as pH and ion concentration require further characterisation. Unlike the simulated liquid environment of the imaging buffer, the cytoplasm in live cells undergoes transformation between glass to liquid forms depending on metabolic activity (Parry et al., 2014). Moreover, the cytosol environment is crowded with all types of proteins and chemicals at distinctive sizes (Spitzer and Poolman, 2013) and undocumented interactions between these proteins and chemicals to BMC might play important role determine the structure and function of BMCs in cells. The isolation of BMC from the cytoplasm in AFM studies enabled better characterisation of surface structures and mechanical properties but limited the detection of in cell modulation and potential cross-talking with cytosolic environment such as interactions with cytoskeletal mobilisation system (Savage et al., 2010).

1.7.4.4 Fluorescent microscopy

The introduction of fluorescent microscopy provides researchers with powerful tools to visualise subcellular activity and decipher the biological functions in living cells, which was unable to be delivered through any other methods described above. In general, through the visualisation of fluorescence emitted by fluorescent proteins (FPs) that introduced to target protein in the live cells, localisation, mobilisation, as well as quantification information of targeted protein (or complex) could be retrieved.

Visualisation of BMCs by fluorescent microscopy has been widely practised in previous studies: For carboxysomes, new carboxysome components CcmP has been verified through the co-localisation with known carboxysomal proteins RbcL (Cai et al., 2013); The carboxysome spatial localisation in cells, as well as the assembly procedure, have been extensively elucidated by time-lapse imaging techniques (Cameron et al., 2013; Savage et al., 2010); Fluorescent microscopy also provides an easy way to validate the assembly of artificial carboxysomal components in native host (Cai et al., 2015b) as well as in non-native species (Fang et al., 2018; Lin et al., 2014a; Sargent et al., 2013). Meanwhile fluorescent microscopy-based techniques such as Fluorescence Recovery After Photobleaching (FRAP) has also been performed on an abnormal bar-like carboxysomes (limited by the FRAP

resolution to regular carboxysome diameter), suggesting the immobilised nature of Rubiscos complexes within formed carboxysomes (Chen et al., 2013); Researchers have also performed fluorescence resonance energy transfer (FRET) to characterise the proximity and formation of CcmK shell proteins within the sheet formation, revealing significant stronger interactions within CcmK2 compared with any other paralog pair (Samborska and Kimber, 2012).

Besides, the *in vivo* quantification is another advantage that has been exploited widely in studies over different cell and organisms (Chong et al., 2015; Hoffman, 2002; Kim et al., 2010; Muzzey and van Oudenaarden, 2009; Novak et al., 2015; Verveer et al., 2000; Zhang et al., 2018). The quantification *in vivo* serves as a good reference to conventional quantification such as immunoassay. Interestingly, the advantages of *in vivo* quantification have never been fully exploited in BMC studies. Signal quantification has only been adopted in few research indicating the event of BMC birth event (Chen et al., 2013; Niederhuber et al., 2017). In the author's view, the differences in tagging strategy are the core reasons limiting the quantification application. In most fluorescent studies, another copy of FP tagged protein is usually inserted in the neutral insertion site (NS) or an inducible expression vector (Cai et al., 2013; Cameron et al., 2013; Savage et al., 2010). In such cases, the wild-type copy of the target protein is maintained. Quantificational fluorescent microscopy can only determine the content of tagged proteins, leaving wild-type proteins un-quantified. In this thesis, we adopted a strategy that replaces the wild-type gene with the FP-tagged version locate in the original genomic loci. In this way, we maintained the native control of wild-type gene and valid the detection of total protein as wild-type gene has been eliminated through segregations through antibiotic selections (described in **Chapter 2**).

As a classical technique that has been developed for decades, on-going development of new fluorescent microscopy variants aiming at better resolution, better precision of quantification is now available. Structured illumination microscopy (SIM) is a new technique with enhanced resolution through computational image reconstruction based on data collected in and outside the normal detection range (Kraus et al., 2017). In BMC studies, the localisations of shell protein CcmK4 and

structural protein CcmM within β-carboxysome *in vivo* are illustrated, therefore further confirmed the shell localisation of CcmK4 (absence in the interior) and even distribution of CcmM within carboxysome lumen (Niederhuber et al., 2017). In addition, special-purposed techniques have also been developed to extend the application of fluorescent microscopy in BMC studies. For instance, a photon-active GFP has been introduced into carboxysomes, suggesting the permeability to protons by the shells, further indicated the absence of pH gradient across carboxysome shells (Menon et al., 2010).

1.8 Aims of this thesis

My PhD research aims at the functional and structural modulation of carboxysomes in native host *Synechococcus elongatus* PCC7942. In **Chapter 3**, we investigate the formation and spatial position of β-carboxysomes through a combination of live-cell confocal fluorescence microscopy and biochemical and physiological approaches. In **Chapter 4**, we further characterise the exact stoichiometry of undocumented building components in individual carboxysome as well as their compositions coping with environmental changes at the single-organelle level using real-time single-molecule fluorescence microscopy, confocal and electron microscopy, combined with a suite of biochemical and genetic assays. In **Chapter 5**, we further document the modulation of carboxysomes under the diurnal light condition and investigated the involvement of circadian clock regulation in the biogenesis and organisation of carboxysomes by live-cell confocal fluorescent microscopy centred techniques.

Chapter 2

Material and methods

Chapter 3

2.1 Bacterial strains and growth conditions

2.1.1 *Escherichia coli*

Escherichia coli (E. coli) strains used and generated in this work were from lab collections. Two *E. coli* strains were used for maintenance, reproduction and manipulation of DNA fragments that were later used in cyanobacteria transformation: DH5α (Thermo-Fisher Scientific) and BW251113 that contain a λ-red recombination system (Datsenko and Wanner, 2000). *E. coli* strains were either cultured in liquid using premade LB broth (Miller's) powder at 25 g/L (Thermo-Fisher Scientific 12795027 Miller's LB broth base) with starting OD_{600} at 0.02-0.05 (Jenway 6300 Spectrophotometer, Jenway, UK) or on LB agar plates made by premixed powder at 37 g/L (Thermo-Fisher Scientific 22700041 Lennox LB agar powder) at 37 ℃. An exception for BW251113 that before recombination procedures the cells were grown at 30°C as the λ-red recombinant plasmid is heat sensitive and inactivation would occur if grown at higher temperature. Cultures were allowed to grow overnight for 8-16 hours where dense cultures reaching OD_{600} at 1.2-1.5 were obtained for plasmid extractions. Antibiotics were supplied at the following concentrations: ampicillin at 100 μg/mL, apramycin at 50 μg/mL and spectinomycin at 50 μg/mL, kanamycin at 50 μg/mL in ddH_2O and chloramphenicol at 10 μg/mL in ethanol. For storage of the strains, overnight cultures were transferred to 4 ℃ for short-term storage and were snap-frozen by liquid nitrogen after mixing with protectant (20% glycerol) and stored at -80 ℃ for long-term storage. All plasmids were also extracted and stored separately at -80 ℃. Strains generated for this work were listed below in **Table 2-1**. Maps for the vectors used in published works are attached in **Appendix B**.

Table 2-1. *E. coli* vectors in this work

Vector names	Description	antibiotic resistance	Related Chapter
pGEM-T easy	Cloning	ampicillin	3,4,5
BW251113	Lambda-red recombination	chloramphenicol	3,4,5
eGFP_pIJ786	Amplification of eGFP: apramycin Cassette	apramycin	3
YFP_pIJ786	Amplification of YFP: apramycin Cassette	apramycin	4,5
GFP free Inducible	Inducible eGFP expressed in Syn7942	apramycin	3
mYpet_pIJ786	Amplification of mYpet: apramycin Cassette	apramycin	4
KO_pIJ788	Amplification of KO: spectinomycin cassette	spectinomycin	5
pLL06_CFP-His-Km_pET26b	Amplification of CFP: kanamycin cassette	kanamycin	4,5
pAM2195	Luciferase reporter with PsbAI promoter	chloramphenicol	5

2.1.2 *Synechococcus elongatus* PCC 7942

Synechococcus elongatus PCC 7942 (hereafter shown as Syn7942) wild-type strain was from lab collections and used to generate mutants in this study. Cultures were grown in BG-11 medium (Rippka et al., 1979) or BG-11 agar plates with TES buffer pH 8.2 (22.9 % w/w of $C_6H_{15}NO_6S$) and sodium thiosulphate (0.3 % w/w of $Na_2S_2O_3$), solidified by 1.5-2 % Agar-agar (w/v). For liquid cultures, final culture volume at 5-10 mL or 30-50 mL were added in small filter capped culture flasks (Nunc™ Cell Culture Treated EasYFlasks™ 156367 Thermo-Fisher Scientific) or large filter capped culture flasks (Nunc™ Cell Culture Treated EasYFlasks™ 156499 Thermo-Fisher Scientific) and kept in 30 ℃ culture room with constant shaking at 100-120 rpm under warm growth light illumination (described in **Figure 2-3**) at 35 $\mu E \cdot m^{-2} s^{-1}$. Strains on plate and in liquid used for inoculation were refreshed every 2-4 and 1 week by dilution with fresh BG-11 medium together with antibiotics, while for mid-term storage cultures were kept at 30 ℃ under 20 $\mu E \cdot m^{-2} s^{-1}$ constant white light illumination. For long-term storage, cell suspensions were washed by fresh BG-11 and concentrated to OD_{750} above 10, then supplemented with 8 % DMSO and immediately snap-frozen by

liquid nitrogen and stored at -80 ℃. The derivative strains produced in this work are listed in **Table 2-2**.

Table 2-2. Cyanobacterial strains generated in this work.

Strain names	Method of generation	antibiotic resistance	Related Chapter	Source/author
RbcL:eGFP	Homologous recombination	apramycin	3	Luning Liu
CcmK4:eGFP	Homologous recombination	apramycin	3	Yaqi Sun/Matthew Faulkner
free GFP	Inducible vector on plasmid	apramycin	3	Yi Fang
CcmK2-YFP	Homologous recombination	apramycin	4	Yaqi Sun
CcmK2-YFP partial		apramycin	4	Yaqi Sun
CcmK3-YFP		apramycin	4	Yaqi Sun
CcmK4-YFP		apramycin	4	Yaqi Sun
CcmL-YFP		apramycin	4	Yaqi Sun
CcmM-YFP		apramycin	4	Yaqi Sun
CcmN-YFP		apramycin	4	Yaqi Sun
RbcL-mYpet		apramycin	4	Yaqi Sun
RbcL-YFP		apramycin	4	Yaqi Sun
RbcS-YFP		apramycin	4	Yaqi Sun
RbcX-YFP		apramycin	4	Yaqi Sun
CcaA-YFP		apramycin	4	Yaqi Sun
CcmP-YFP		apramycin	4	Yaqi Sun
CcmP-YFP partial		apramycin	4	Yaqi Sun
CcmO-YFP		apramycin	4	Yaqi Sun
CcmO-YFP & RbcL-CFP		apramycin & kanamycin	4	Fang Huang
KaiA-YFP		apramycin	5	Yaqi Sun
ΔKaiA		spectinomycin	5	Yaqi Sun
ΔKaiA & RbcL-YFP		apramycin & spectinomycin	5	Yaqi Sun
KaiA-YFP & RbcL-CFP		apramycin & kanamycin	5	Yaqi Sun
pAM2195	Neutral Site insertion	chloramphenicol	5	Plasmid gift from Golden lab

2.2 Molecular biology

2.2.1 DNA transformation of *E. coli*

2.2.1.1 *E. coli* Component cell preparation

Component cells used in this work were prepared freshly with DH5α and BW251113 cells (see **section 2.1.1**). For DH5α cells, 2 μL of overnight culture or half a single colony from agar plate were inoculated with 10 mL liquid LB medium for overnight growth at 37 ℃. On the day of the heat shock

transformation, 100 μL of the overnight culture were added into 40 mL fresh LB medium (separate in 4 vials for sufficient mixing) and grown for 3 to 4 hours at 37 ℃ until density reaching OD_{700} between 0.5-0.6. Cells were then washed with 5 mL of ice water pre-cooled $MgCl_2$ solution (100 mM, autoclaved) for three times by centrifuging at 4500 g, 4 ℃ for 3 minutes. The final pellet was resuspended with 1 mL of ice-cold $CaCl_2$ solution (100 mM, autoclaved) and kept on ice for at least 10 minutes before use.

For BW251113 cells, on the day of the electroporation, cultures were inoculated from 2 μL of overnight culture or half a single colony from agar plate into 10 mL liquid LB medium for overnight growth at 30 ℃, re-inoculated from 100 μL into 30 mL of fresh LB medium with chloramphenicol for 4 to 4.5 hours until density reach 0.5 to 0.6, measured by spectrophotometer at OD_{700}. Then similar wash steps were involved but with ice-cold 10 % glycerol solution (autoclaved) for three times. The final pellet was resuspended in 100 μL of 10 % glycerol solution (autoclaved) and kept on ice for at least 10 minutes before use.

2.2.1.2 Heat shock

Heat shock was carried out to integrate the gene fragments together with its 1.5 kilobase pair (kbp) upstream/downstream flanking sequence onto a pGEM-T easy vector according to protocol described (Froger and Hall, 2007) with several modifications. 500 μg of extracted DNA fragments or 5 μg plasmid were added to 100 μL of chemically competent *E. coli* cells and incubated on ice for 20 minutes. The cells were heated shocked for 45 s at 42 ℃ then placed back immediately on ice and kept on ice for 2 minutes. Then 1 mL of ice-cold LB medium was added to the mixture without vortexing, sealed with foil and place on the shaker in 37 ℃ growth room for at least 1 hour. After the growth, 200 μL of the mixture was spreaded onto plates supplemented with Isopropyl β-D-1-thiogalactopyranoside (IPTG), 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) and antibiotics. Negative control was done in parallel with just component cell. After overnight growth at 37 ℃, blue-white screening was carried out and positive colonies were transferred to new plates for storage and plasmid extractions.

2.2.1.3 Electroporation

We performed two steps of electroporation to introduce the gene fragments including up/down-stream flanking sequences, and induce the recombination by λ-red recombination system (Datsenko and Wanner, 2000) respectively.

During the 1st electroporation, 50 μL of BW251113 competent cells (described in **section 2.2.1.1**) were mixed with 5 μL of p-GEM plasmid ligation product that contain target gene fragments at manufacturer suggested concentration, placed in pre-cooled 2 mm-gap electroporation cuvette for 5 minutes. The electroporation was performed with 2500 V, 25 μF of capacitance, 200 Ω of resistance by Bio-Rad electroporation system (Bio-Rad Gene Pulser equipped with a pulser controller and a capacitance extender module). After the pulse, 1 mL of ice-cold LB medium was immediately added in and mixed well by pipetting, then transfer back to 1.5 mL EP tubes for 1-1.5 hours of growth in 30 ℃ culture room for cell multiplication. Later the cell culture was pelleted down by centrifugation at 13,000 rpm for 1 minute. 950 μL of supernatant was removed, leaving the pellet resuspended with the remaining 100 μL LB medium later to be spreaded onto a plate with ampicillin and chloramphenicol and grown overnight at 30 ℃. Colony screening PCR was performed (described in **section 2.2.4**) to confirm the successful transformation.

The 2nd electroporation was performed similarly to the 1st one. The differences are: firstly, after the pulse, ice-cold liquid LB supplemented with 200 mM of arabinose was added instead of just LB to the cells to induce the production of λ-red recombinase. Secondly, the growth duration after electroporation was increased to 2-2.5 hours to allow sufficient recombination. Thirdly, the overnight growth was done with antibiotics on the successful recombinant at 37 ℃ instead of 30 ℃. Moreover, lastly, the colony screening PCR was performed using colonies from the 2nd or 3rd plates instead ones from the 1st plate to allow duplications of recombinant plasmids with present of respective antibiotics and eliminations of previous non-recombinant plasmids through removal of antibiotic pressures.

Colony screening PCR was performed with segregation primers (listed in **Appendix A**) to confirm the successful recombination.

2.2.2 DNA Transformation of Syn7942

All fluorescent protein tagging and Knock-Out mutants were generated following the same redirect strategy (Liu et al., 2012). For fluorescent protein insertion, cassettes containing coding sequence of target protein plus fluorescent protein on its C-terminus, as well as an antibiotic-resistant protein operon could replace the Wild Type copies of the target gene on original genome through homologous recombination. For Knock-Out mutants, the Wild Type copies of the target gene were replaced by cassettes containing only antibiotic-resistant genes, therefore eliminate the target gene on genomes through recombination. The diagram of recombination is shown in **Figure 2-1**.

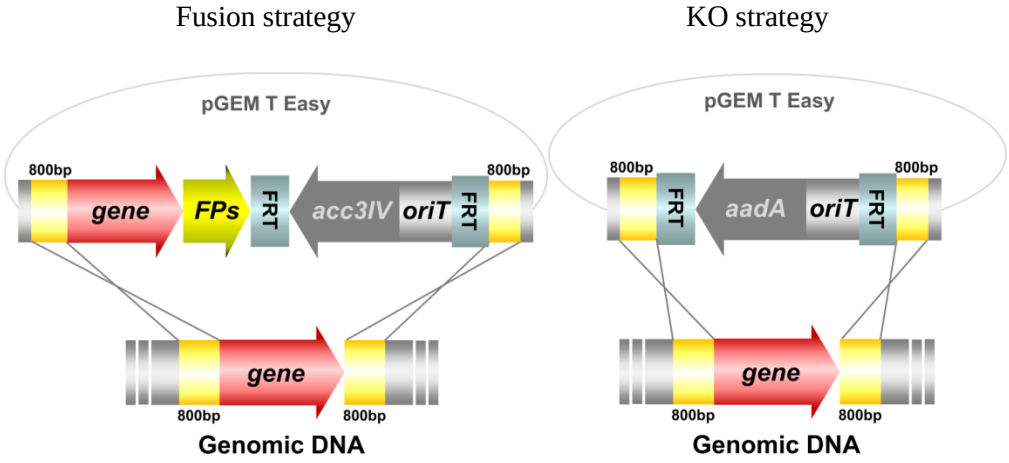


Figure 2-1. The strategy of FP fusion and Knock-Out using REDIRECT protocol. A fragment contains 800 bp upstream and downstream of the gene with or without gene plus fluorescent protein together with an antibiotic-resistant cassette (encoded in reverse orientation) in order to replace the genomic DNA fragment by homologous recombination. FRT indicates flippase recognition target which can be used to excise inserted cassette. *Acc3IV* and *aadA* indicate apramycin and spectinomycin resistant genes respectively. OriT indicates short sequence as the origin of transfer during bacterial conjugation.

2.2.2.1 Syn7942 Component cell preparation

Syn7942 cells are naturally component for transformation and efficiency for acceptance of foreign DNA depends on the status of growth (Golden and Sherman, 1984). Therefore in this thesis

component cells were prepared freshly. 40 mL cultures in exponential growth phase were harvested by centrifugation (5,000 g, 30 ℃ for 10 minutes) and washed with BG-11 medium three times and then concentrated to 1 mL. The dense component cells were kept at 30 ℃ before use within 30 minutes.

2.2.2.2 Homologous recombination of Syn7942

Syn7942 was transformed with the recombinant plasmid DNA extracted from *E. coli* cells after 2nd electroporation as described earlier (Liu et al., 2012). For each transformation, 30 μL of recombinant plasmid was added to 100 μL of cells and incubated for overnight at 30 ℃ or 3 hours at 34 ℃ in dark. The mixture was then spreaded onto BG-11 agar plates with corresponding antibiotics. Cells on plates were kept grown under maintenance condition (described in section **2.1.2**) for 2-3 weeks. The surviving colonies were transferred to new plates for further growth and screening.

2.2.2.3 Evaluation of homologous recombination in Syn7942 mutants

Due to the existence of multiple copies of genome in Syn7942, the recombination status needs to be further checked by colony screening (described in **section 2.2.4.1**). The evaluation of segregation was performed with primers designed across the insertion/deletion site, where different sizes of band indicate non-insert or successful-insertion/deletion respectively, as shown in **Figure 2-2.** Primers used during screening were all listed in **Appendix A**. During screening, a negative control using Wild Type genomic DNA as template was also included. The sizes of amplified bands were estimated from DNA ladder that loaded together with the samples during gel electroporation.

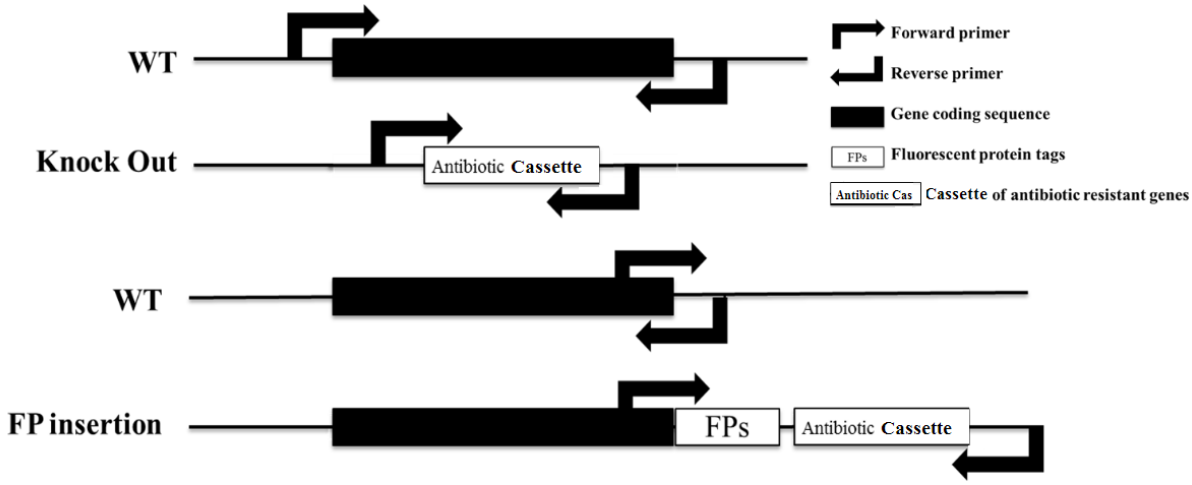


Figure 2-2. Schematic diagram of segregation screening. Primers were designed over the up/ down-stream of gene for Knock-Out, where successful recombination would yield fragments at size equal to the size of Wild-Type band, minus the size of the coding sequence, then plus the size of the spectinomycin resistant fragments at 1412 bp. For fluorescent protein insertions, primers were designed before and after the stop codon of the target gene, leaving the fragment amplified with Wild-Type template at averagely 200 bp while successful recombination results in fragment sizes that are 2126 bp greater than their Wild Type counterpart.

2.2.3 DNA purification and isolation

2.2.3.1 Genomic DNA extraction from Syn7942 cells

Genomic DNA from Syn7942 was extracted with PureLink Genomic DNA Mini Kit (Fisher Scientific UK Ltd) following the manufacturer protocol. The genomic DNA extracted from Wild-Type Syn7942 was used as template for carboxysome gene and fragment amplification via PCR (described in section **2.2.4**). Extracted DNA was then checked for concentration and quality (described in section **2.2.3**) before storage.

2.2.3.2 Plasmid purification from *E. coli* cells

Plasmid extraction from *E. coli* was done using the GeneJET Plasmid Miniprep kit (Fisher Scientific UK Ltd) according to the manufacturer's instructions. The extracted plasmids were then checked for concentration and quality before stored at -20 ℃ for short-term and -80 ℃ for long-term storage.

2.2.3.3 PCR product purification after gel electroporation

Band of PCR product in agarose gel was cut out after electroporation for purifications by GeneJET Gel Extraction Kit (Fisher Scientific UK Ltd) following the manufacturer's instruction. DNA products were stained by Midori green Nucleic staining solution. To minimise degradations to the DNA, visualisation of agarose gel was done with long wavelength UV at controlled durations. The extracted DNA fragments were then checked for concentration and quality (**section 2.2.3**) before stored at -20 ℃ for short-term storage and -80 ℃ for long-term storage.

2.2.3.4 DNA quality and concentration measurement

The concentration and quality of the DNA were checked by a Nanodrop (ND-100 Spectrophotometer) at Absorbance wavelength of 260 nm (A_{260}) and 280nm (A_{280}). The purity of DNA could be indicated by the A_{260}/A_{280} ratio ranging between 1.85-1.95, where values outside that range would indicate the presence of contaminants such as protein and ethanol from the extraction process.

2.2.4 DNA manipulation

2.2.4.1 Polymerase Chain Reaction

General PCR was done according to the following steps. For each primer set, an initial gradient PCR was performed with template from positive control to determine the optimised annealing temperature ranging from 45-65 ℃ with protocol as follows: Initialisation at 94 ℃ for 2 minutes > (denaturation at 94 ℃ for 45 seconds > annealing at 45-65 ℃ for 45 seconds > elongation at 72 ℃ for 2 minutes) x 15 cycles > (denaturation at 94 ℃ for 45 seconds > annealing at 55 ℃ for 45 seconds > elongation at 72 ℃ for 2 minutes) x 15 cycles > final elongation at 72 ℃ for 4 minutes. The most effective annealing temperature could be determined corresponding to the strongest amplification during gradient PCR. Optimised annealing temperature was used to perform PCR for two purposes. One is to precisely amplify the fragments from genomic DNA, plasmid for manipulation; the other is for quick screening of colonies and fragments. The complete list of primers used in this thesis were list in **Appendix A**.

2.2.4.1.1 Non-high fidelity PCR for screening purposes

The DreamTaq polymerase (Thermo Fisher Scientific) was generally used for non-high fidelity amplification of DNA during screening of colonies for *E. coli* and Syn7942 strains. Standard protocol was used according to manufacturer manual. An additional procedure was performed for template preparation from Syn7942 cells before used as template for PCR reactions, where cells were harvested and washed with ddH_2O and denatured at 99 ℃ for 5 minutes. This step is to ensure the successful release of DNA from the cell and minimise the potential effect brought by the salt-enriched BG-11 medium during PCR reactions.

2.2.4.1.2 High fidelity PCR

High fidelity PCR was performed to amplify DNA precisely. In this work, Q5 Hot Start High-Fidelity DNA polymerase (New England Biolabs) and CloneAmp HiFi PCR Premix (Clontech) were used following the manufacturer protocols. An additional A-tailing procedure by DreamTaq polymerase was done over gene fragments for ligation purposes to generate blunt poly-A end region that was required by pGEM-T easy vector ligation.

2.2.4.2 DNA sequencing

To confirm the final constructs generated are correct in sequence, sequencing was performed using either extracted, genomic DNA or high fidelity amplified PCR product. The sequencing service was provided by Eurofins, with full sequencing coverage from both strands.

2.3 Treatments on Syn7942 strains

2.3.1 Light and CO_2 treatment conditions

For light treatments, the light treatments in chapter 3 for light modulate biogenesis study and chapter 5 for stoichiometry study are slightly different. *Synechococcus* cultures were provided with light at intensities of 10 $\mu E \cdot m^{-2}\ s^{-1}$ as LL, 50 $\mu E \cdot m^{-2}\ s^{-1}$ as ML, and 100 $\mu E \cdot m^{-2}\ s^{-1}$ as HL respectively in **Chapter 3**. In **Chapter 4**, light intensity of HL treatment was reduced to 80 $\mu E \cdot m^{-2}\ s^{-1}$ due to the

difficulties of carboxysome signal quantification at 100 $\mu E \cdot m^{-2}$ s^{-1}. With reduced light intensity but still higher than ML, the HL treatment in **Chapter 4** was capable of representing the phenotype induced by increased light meanwhile maintained separated signal profiles for more effective quantification analysis.

The CO_2 treatments were done in a specially designed incubator (SciQuip Incu-430C) with light illumination at 50 $\mu E \cdot m^{-2}$ s^{-1}, supported by a CO_2 control module (built upon WEST P6100 controller) that monitor and maintain the CO_2 level at 3% in the chamber. The spectrum of light treatments and CO_2 treatments were shown in **Figure 2-3**.

For all treatments, cultures were inoculated from seeding cultures during exponential growth phase under maintenance condition by dilution of fresh BG-11 medium, kept on same shakers under same rpm settings (100-120rpm). Culture volumes in flasks were controlled as well to ensure sufficient fixing. For large and small culture flasks (Thermo-fisher brand), 40mL and 5mL cultures at final volume were set for all treatments.

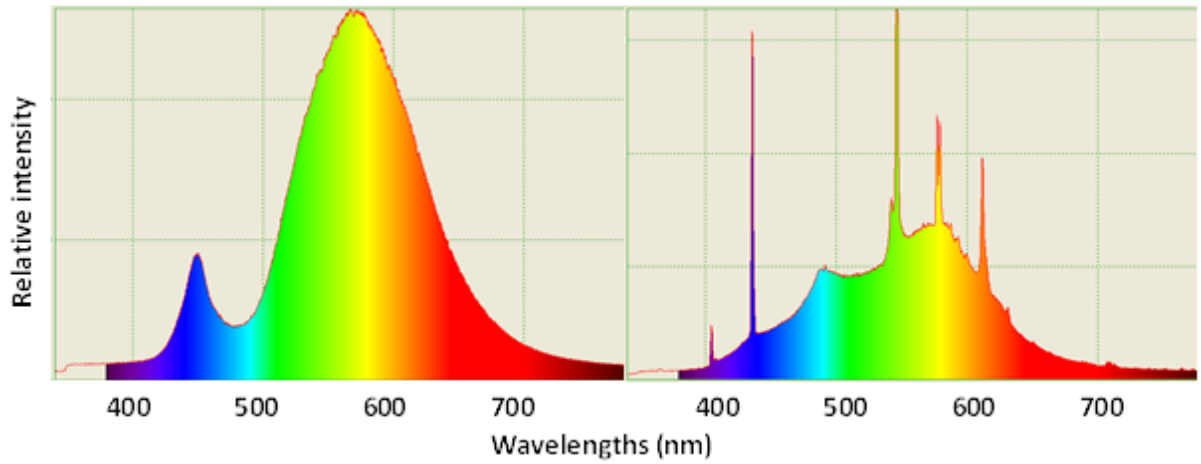


Figure 2-3. Spectrum for light illumination during light treatments and CO_2 treatment. Light quality profiles are similar and comparable between light (left figure) and CO_2 (right figure) treatments, were major peaks appear at 575nm and the 2nd peak at 450nm, representing warm and cold wavelength of light.

2.3.2 Diurnal light treatments conditions

Diurnal light treatments were done according to method described in (Cohen et al., 2014). 12 hours dark-12 hours light were provided on a platform with black curtains that block all surrounding lights. Light intensity was kept same with constant ML treatment at 50 $\mu E \cdot m^{-2}\ s^{-1}$. Seeding cultures were maintained in constant light condition as described in **section 2.1.2**.

2.3.3 Electron transport inhibitor, general protein synthesis inhibitor and induction treatments.

The electron transport inhibitors DCMU (Sigma-Aldrich) and DBMIB (Sigma-Aldrich), and the protein synthesis inhibitor lincomycin were added to 20 μM, 10 μM, and 400 $\mu g \cdot mL^{-1}$, respectively. Cells were adapted for 24 h in the presence of DCMU, DBMIB, or lincomycin before microscopy imaging. The free-eGFP-expressing *Synechococcus* cells were induced with 1 mM of isopropyl β-d-1-thiogalactopyranoside (IPTG) for 24 h before confocal imaging.

2.4 Syn7942 strain evaluations

2.4.1 Optical measurement for liquid culture cell densities, absorption spectra and chlorophyll a concentration

For cell density recording and doubling time calculation, the cell densities of culture were tracked through measuring of 750 nm absorption per 24 hours. For larger culture, samples were taken and loaded in 10 mm vials at volume of 1 mL and measured by spectrophotometer (Jenway 6300, UK). For small-scale cultures, samples were taken and loaded in 96-well microplate at volume of 100 μL and measured by microplate reader (Spectramax 340, Molecular Devices, US), read at wavelength 750 nm.

For absorption spectra measurement, samples were taken and loaded in 96-well microplate, then measured by microplate reader at wavelength from 500 nm to 750 nm. For chlorophyll a (Chla) content determination in oxygen uptake measurement of the strains, standard protocol was applied as

previously described (Komenda and Barber, 1995). Cell cultures were washed and concentrated 10 times (final OD_{750} ranging from 5 to 20). Then 10 μL of concentrated cell cultures were mixed well with 990μL of methanol by pipetting and let settled for 10 minutes. The mixture was then centrifuged at 12,000 g for 2 minutes by bench-top centrifuge for pellet removal). The supernatant was collected and transferred in a 1 mL cuvette for absorption measurements at both 666 nm and 750 nm by Spectrophotometer. Chl a concentration (μg/mL) was calculated by the following equation:

$$C_{(\text{Chl a})} = (OD_{666} - OD_{750}) \times 12.6 \times 10$$

In the equation, 12.6 is the Chl a molar extinction coefficient and 10 is the dilution factor. Standardisation of the spectrophotometer was done using mixture made with 10 μL fresh BG-11 and 990 μL methanol.

2.4.2 Oxygen evolution measurement

O_2 uptake of cell cultures was measured in the dark at 30 ℃ in a Clarke type oxygen electrode (OxyLab 2; Hansatech) according to method described in (Liu et al., 2012). In general, 1 mL of cell suspension calibrated to Chl a concentration of 10 μM was added into the electrode chamber under constant temperature control at 30 ℃, sealed from the atmosphere with sufficient mixing. The readings were taken after the culture has been adapted in chamber in darkness for 2 minutes. Meanwhile the oxygen consumption rate was recorded. Light module was then turned on at maximum capacity to record the oxygen evolution rate. For each biological sample, at least three technical repeats were done to ensure accurate measurement. For each strain, four biological repeats were measured. Preparation of oxygen electrode and membrane was performed following manufacturer's manual.

2.4.3 Carbon fixation efficiency measurement

For CO_2 fixation capacity measurement that solely depending on the quantity and functionality of carboxysomes, cells were harvested at exponential phase under corresponding light treatments and then resuspended in Rubisco assay buffer (100 mM EPPS buffer, pH 8.0, and 20 mM $MgCl_2$). Cell density was then calibrated by measuring OD_{750} (described in section **2.4.1**). Radiometric assay was carried out according to a previously described protocol (Price and Badger, 1989) with additional cell permeabilisation treatment (Schwarz et al., 1995). The duration of permeabilisation was tuned as shown in **Figure 2-4**, where 1 minute of reaction time was selected as standard protocol.

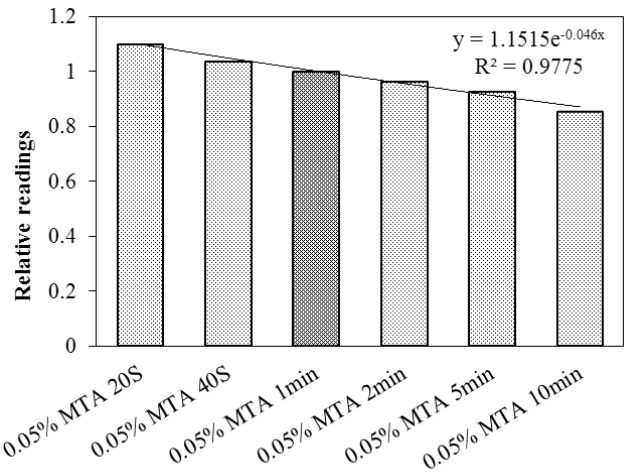


Figure 2-4. Relative count readings from whole cell under different durations of MTA treatment. Longer duration of treatment reduced the detected count readings, the reducing can be linearly fitted, as indicated in the figure.

Cell cultures prepared in assay buffer with the same cell density were incubated with $NaH^{14}CO_3$ (final concentration at 25 mM) at 30 ℃ for 2 minutes and then permeabilised by mixed alkyltrimethylammonium bromide (MTA, final concentration at 0.03 % [w/v]; Sigma-Aldrich). RuBP (Sigma-Aldrich) was then added with a range of concentrations (0–2.0 mM) to initialise the fixation. After 5 min, 10 % formic acid was added to terminate the reaction. Samples were then dried on heat blocks at 95 °C to remove unfixed $NaH^{14}CO_3$, and the pellets were resuspended in distilled water in the presence of scintillation cocktail (Ultima Gold XR; Perkin-

Elmer). Radioactivity measurements were carried out using a scintillation counter (Tri-Carb; Perkin-Elmer). Raw readings were processed to determine the amount of fixed ^{14}C, calibrated by blank cell samples without providing RuBP, and then converted to the total carbon fixation rates. Carbon fixation rates of cell cultures were normalised depending on the AtpB quantity from immunoblot analysis or cell density measured at OD_{750} depending on the purposes of comparison. AtpB is for calibration of fixation rate to equal cellular metabolic levels (used in **Chapter 3**), while total protein concentration is to enable comparison with cell fractions (lysate or isolated carboxysome). The cell density is for quicker and easier comparison of strains (adopted in **Chapter 4** and **5**) based on the assumption that cellular metabolic levels are comparable within strains, indicated from the result obtained in **Chapter 3**. For each experiment, at least three biological repeats were prepared. Significance was assessed by a two-tailed *t*-test.

For *in vivo* carbon fixation rate measurements that were determined not solely by carboxysome quantity and functionality, *in vivo* assays were carried out by providing 2 mM of radioactive sodium bicarbonate ($NaH^{14}CO_3$) into the BG-11 growth medium in small culture flasks (Nunc™ Cell Culture Treated EasYFlasks™ 156367 Thermo-Fisher Scientific) with caps sealed to minimise gas exchange from atmosphere during assay. The cultures were then placed back into respective light treatments for 30 minutes of growth. Later, culture was mixed well, then taken by pipetting at volume of 500 µL into EP tubes with pre-added 200 µL 10 % formic acid. The mixture was then mixed immediately and placed on heat blocks at 95 ℃ overnight to remove unfixed $NaH^{14}CO_3$. Follow on steps were carried out same to ones for the maximum carbon fixation capacity measurement described in previous paragraph.

2.4.4 SDS-PAGE and Immunoblotting

Syn7942 cell samples for immunoblot analysis by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) were washed by TE buffer with Protease Inhibitor Cocktail (PIC) at 1 % (v/v). Then homogenised by sonication (5 seconds sonication with 15 seconds cooling on ice) for 24 cycles at 4 ℃, followed by 2 % (w/v) Triton treatment in dark for 1 hour. Then the lysates were

centrifuged at 4000g for 5 minutes to get rid of the cell debris and unbroken cells. Supernatant collected was then measured for protein concentration by Pierce Coomassie (Bradford) Protein Assay Kit (Thermo Fisher Scientific), diluted to same concentration and then mixed with 4 x SDS-PAGE sample loading buffer (250 mM Tris–pH 6.8, 8% (w/v) SDS, 0.2% (w/v) bromophenol blue, 40% (v/v) glycerol, 20% (v/v) β-mercaptoethanol). Samples were then denatured at 99 ℃ for 10 minutes, vortex for 1 minute for homogenising, then centrifuged at 14,000 g for 2 minutes to get rid of the insoluble fractions. Supernatants were kept at 4 ℃ before loading and store at -20 ℃ for storage.

Supernatants at final loading amount of 75 μg/well (For visualisation of YFPs in **Chapter 4**, loading amount was increased to 150 μg/well for detection of the minor proteins) were loaded on 10% (v/v) denaturing SDS-PAGE gels with PageRuler^{TM} plus Prestained Protein Ladder. PVDF membrane was activated by 100 % methanol for 1 minute and then kept in transfer buffer (25 mM Tris base, 150 mM glycine, 10% methanol) for at least 5 minutes before use. Gels were then electroblotted onto a PVDF membrane (Bio-Rad, US) by wet transfer protocol at 90 V for 45 minutes in pre-cooled transfer buffer at 4 ℃. The membrane was collected and soaked in TBS buffer (50 mM Tris pH 7.5, 150 mM NaCl) before blocking by 5 % (w/v) skim milk in TBS for 2 hours at room temperature or overnight at 4 ℃. Membrane was then washed with TBST buffer (50 mM Tris pH 7.5, 150 mM NaCl, 0.1 % (v/v) Tween 20) for 10 minutes, then incubated with primary antibodies of mouse monoclonal anti-GFP (Life Technologies, UK) for both eGFP and YFP recognition, rabbit polyclonal anti-RbcL (Agrisera, Sweden) for RbcL recognition, anti-ATPaseB (Agrisera, Sweden) antibodies for AtpB for 4 hours in room temperature or 4 ℃ overnight on a swing shaker. Additional washes by TBST were done afterwards for 4 x 10 minutes before incubation with second antibodies of horseradish peroxidase-conjugated goat anti-mouse immunoglobulin G secondary antibody (Promega, US) or anti-rabbit immunoglobulin G secondary antibody (GE Healthcare, US) for 2 hours at room temperature. After incubation, the membrane was further washed with TBST for 4 x 10 minutes and kept in TBS before visualisation.

Signals were visualised using the Clarity Western ECL Substrate Kits (Bio-Rad, US) and captured by ImageQuant LAS 4000 (GE Healthcare Life Sciences). AtpB protein was used as a loading control for cell population (Zhang et al., 2012) in **Chapter 3**. In **Chapter 4** total protein concentration was used as loading control instead to enable comparison with previous immunoblot data (Long et al., 2010). Immunoblot protein quantification was carried out using ImageJ. For each experiment, at least three biological repeats were performed.

2.4.5 Luciferase reporting system for circadian detection

Cells containing luciferase reporter were pipetted on the BG-11 plate as a droplet and grown in diurnal light treatments for four days before imaging. Then the plate was placed in a light-tight imaging box for bioluminescence capturing for 10 minutes by ImageQuant LAS 4000 (GE Healthcare Life Sciences) using high-sensitive protocol. Then the plate was placed back to diurnal light treatments until next imaging in a 2-hour imaging interval. Single intensities for the entire cell patch was analysed by ImageJ for quantificational comparisons. The working method was adapted from protocol as described for strains that express pAM2195 which is autonomously bioluminescent (Mackey et al., 2007).

2.4.6 TEM imaging and analysis

Thin-section Electron Microscopy (TEM) for whole cell Syn7942 were done from cultures grown for light and gas treatments that were also used for confocal imaging. 20mL of cultures were provided and washed by EM buffer (0.05 M sodium cacodylate, pH = 7.2) for three times, then fixed in 2 % (v/v) glutaraldehyde and 2 % (v/v) paraformaldehyde in EM buffer for 1 hour. Then cells were post-fixed by 1 % (v/v) osmium tetroxide for 1.5 hours and dehydrated with a series of increasing alcohol concentrations. After dehydration, stepwise embedment was done over transmit EM resin (TAAB Laboratories Equipment, UK) (acetone/Transmit (1:1) for 20 minutes, and acetone/Transmit (1:2) for 4 hours. Samples were then incubated for additional 6h and left to polymerise for two days at 70 ℃ Thin sections of 70 nm were cut with a diamond knife (Diatome 45 ° ultra, Agar Scientific) and post-stained with 4 % uranyl acetate and 3% lead citrate. Images were recorded by FEI Tecnai G2 Spirit

BioTWIN transmission electron microscope. The diameter of carboxysome was measured from TEM images manually according to the method described in (Faulkner et al., 2017). To reduce the deviation caused by sections close to the out rims of carboxysome, only carboxysomes with typical polyhedron shapes were kept in the sample pool. Diameter data was analysed by Origin software.

2.5 Fluorescence microscopy imaging

2.5.1 Confocal Imaging sample preparation

Sample preparation was done as described earlier (Liu et al., 2012). In general, Syn7942 cells that were grown in liquid culture from the culture flasks were taken by pipetting at 10 μL and applied over 1.5 % agar (w/v) BG-11 plate with corresponding antibiotics, leaving to dry in fume food for 10 minutes. The agar with cell patches was then cut out and applied on a 0.17 mm glass coverslip, where cells were sandwiched between agar and glass cover therefore immobilised for imaging. The BG-11 plates used were pre-incubated at 30 ℃ before using to minimise temperature alternation for the cells.

2.5.2 DAPI staining

DAPI (4',6-diamidino-2-phenylindole) staining was done as described (Smith and Williams, 2006). Cells were collected after four days of different light treatments by centrifugation at 6,000 g, 30 ℃ and were then washed by PBS (pH=7.2) for three times and stained by DAPI solution in dark for 20 minutes, 30 ℃. After staining, cells were washed by water twice before giving drops onto a BG-11 plate for microscopy. The DAPI-stained DNA contents were visualised by Zeiss LSM780 excited by 303 nm laser. The emissions were recorded between wavelengths between 430 nm and 480 nm.

2.5.3 Confocal imaging

For non-quantificational imaging, laser-scanning confocal microscopy used a Zeiss LSM510 or LSM710 with 63 × or 100 × oil-immersion objectives and excitation at 488 nm by argon laser, under a pinhole setting at 1 μm. Images were captured using Zeiss Zen 2010 software. GFP and chlorophyll fluorescence were detected between 500-520 nm and 660-700 nm respectively. Transmission images

were also captured when possible. Live-cell images were recorded from at least five different cultures. All images were captured with all pixels below saturation.

For quantificational imaging, confocal laser scanning confocal microscopy used a Zeiss LSM780 with alpha Plan-Fluar 100 x /1.45 Oil objective and excitation at 514 nm from Argon laser for relative signal quantification. YFP signals were captured within emission between 520-550 nm. Chlorophyll auto-fluorescence signals were captured within emission between 660-700 nm, recorded as 512 x 512 pixels images at 16 bits. Pinhole was set to 1μm to ensure full capture of carboxysomal signals. Live-cell images were recorded from at least five different cultures. All images were captured with all pixels below saturation. The sample platform was pre-incubated and thermo-controlled at 30 °C before and during imaging. Zoom settings were set to have each carboxysome visualised with a minimum of 8 x 8 pixels array to allow sufficient profiling of carboxysomal signals by peak intensity recognition and measurement. For CcmP-YFP with minimum signal intensity, visualisation was done by LSM880 with Airyscan mode (Huff, 2015). CcmO-YFP/RbcL-CFP dual fluorescence imaging was done with 458/514 dual lasers that excite both CFP and YFP. Emissions were then captured within wavelengths at 520-550 nm and 470-500 nm for YFP and CFP respectively.

Furthermore, the argon laser at 25% power was warmed up for more than 1 hour to achieve stable power output. The sample platform was pre-incubated and thermo-controlled at 30 ℃ before and during imaging. Laser power and gain settings (laser power, gain, offset, zoom, pixel size, depth) were determined and maintained same to have good SNR while minimising signal saturation for all strains and conditions in groups of comparison. Focus was found by scanning on adjunctive region beside ROI. Therefore bleaching by the laser during scanning was avoided. Imaging was done within 30 minutes after harvesting the cell from treatment platform to minimise the adaptations to environmental changes.

2.5.4 Time-lapse imaging

Sample preparation was modified based on protocol suggested in (Yokoo et al., 2015). 2 mm thin Agar mat with BG-11 was prepared in stacked sandwiches to contain drops of diluted cells. Cell was allowed to grow on agar mat for 1-2 hours before imaging under provided light over the microscope. A switch module that controls the light on/off was connected to disable light illumination during laser scanning. The laser power was set to minimum (1 %) to reduce the bleaching for signals during long-term tracking. Cells were incubated on the BG-11 agar mat on the microscope for 1-2 hours before imaging The continuous light illumination was provided at intensity calibrated equally to HL/ML/LL for medium culture growth which maintain the respective carboxysome count per cell confirmed through on-scope pre-growth. The same illumination was applied to the cells during time-lapse imaging with a hand-made module that switched off the light during laser scanning (less than 5 s per minute intervals). The interval time was set to 60 s to guarantee sufficient light illumination between imaging. The laser power was set to the minimum (1%) to reduce the bleaching for signals during long-term tracking.

2.5.5 Slimfield imaging

Cells were applied at the small volume onto the BG-11 agarose pad at 0.25 mm thickness, air dried and then assembled with glass coverslips, plasma cleaned (Harrick-Plasma) for ~1 min. A dual-colour bespoke laser excitation single-molecule fluorescence microscope was used utilising narrow epifluorescence excitation of 10 mm full width at half maximum (FWHM) in the sample plane to generate Slimfield illumination (Wollman et al., 2017). This was incident on a sample mounted on a Mad City Labs nanostage built on an inverted Zeiss microscope body consisting of a 20 mW 514 nm laser. A Chroma GFP/mCherry dichroic was mounted under the Olympus 100x NA = 1.49 TIRF (total internal reflection fluorescence) objective, which delivers 10 mW excitation power. The image was split into YFP and chlorophyll channels using a bespoke colour splitter utilising a chroma dichroic split at 560 nm with 542 nm and 600 nm, 25 nm bandwidth filters. Imaging was done with an Andor iXon 128 EMCCD camera (iXon DV860-BI, Andor Technology, UK), at a pixel magnification

of 80 nm/pixel using 5 ms camera exposure time. Excitation intensity was initially reduced by 100x using and ND=2 or 1 filter for high copy number strains (all except CcmL and RbcX) to avoid pixel saturation on the EMCCD camera detector before a full-power photobleaching.

2.6 Microscopy data analysis

General image analysis was carried out using ImageJ software (NIH Image, Bethesda, US) and Fiji (Schindelin et al., 2012) for intensity profiling, carboxysome recognition. Image SXM (www.ImageSXM.org.uk) was used for statistical analysis of carboxysome numbers per cell, carboxysome distribution within cells, as well as dimensions of cell length/width measurements.

2.6.1 Carboxysome recognition

Image SXM was used for carboxysome recognition in **Chapter 3** in RbcL-eGFP and CcmK4-eGFP strains. The recognition algorithm is optimised by the author of the software for these two strains. Besides, in **Chapter 3** and **4**, lower SNR (Signal to Noise Ratio) images for minor carboxysomal proteins with weaker signals, ImageJ and Fiji Plugin "Find maxima" plugin-based local maxima detection algorithm was adopted instead (https://imagej.nih.gov/ij/docs/guide/146-29.html#toc-Subsection-29.4). Noise tolerance was determined by background intensities in empty regions. Imaging for different treatments in the same strain was performed under the same imaging settings. For strains with obvious cytosolic signals, peak intensities were further subtracted for cytosolic backgrounds, which were determined by the average peak intensities in non-carboxysome regions over the central line of the cell. Raw data was processed by Origin Lab and MATLAB (Mathworks) for profile extraction and statistical analysis and the goodness-of-fit parameter for Violin plot visualization. Violin plots were generated by R to illustrate the fluorescence intensity distribution of individual building proteins per carboxysome fitted by kernel smooth fitting. The representative values and deviations of signal intensities were represented by Peak value ± half width at half maximum (HWHM) measured from kernel density fitted profiles, respectively. The significance of differences between treatments were evaluated by Man-Whitney U-tests pair-wisely. Standard errors

of sampling was determined through randomized grouping of intensity entries, with each group containing a minimum of 70-100 entries. Errors were controlled below 5% to have accurate estimation from the distributions.. A typical recognition overlay is shown below as **Figure 2-5**.

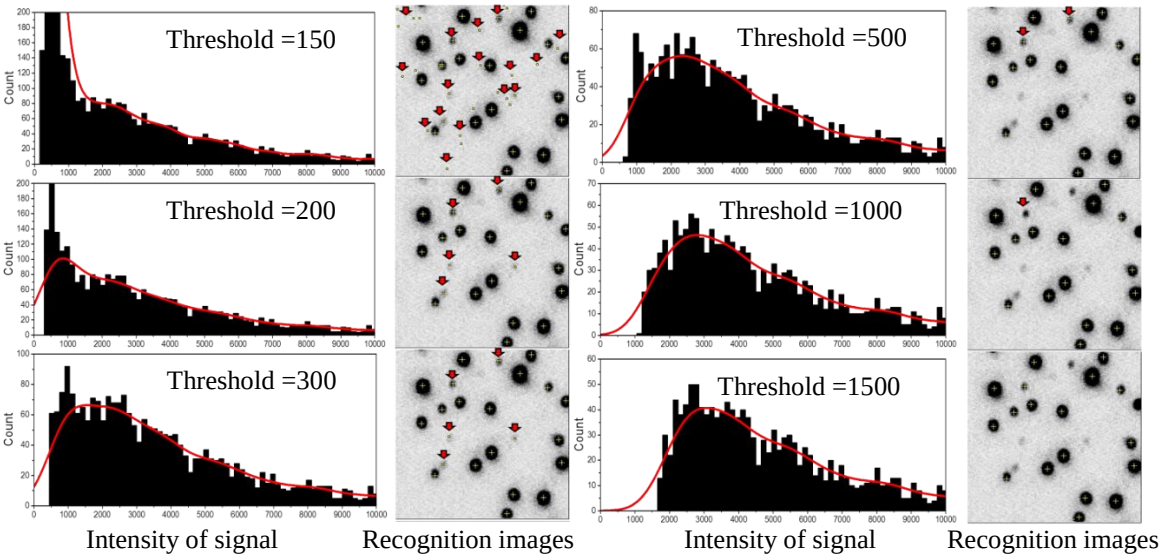


Figure 2-5. Typical carboxysomal signal recognition by ImageJ "Find maxima" method under different noise tolerance settings from 150, 200, 300, 500, 1000, 1500 for RbcL-YFP. Signal histogram and density plot overlays were shown on the left column while corresponding confocal images with channel for carboxysomal signals, brightened and contrast reduced for weaker signal illustration were shown on the right. The background intensities measured from no cell region is 489 ± 227, n = 5 as selected views. The final threshold was then set to 500. The representative intensity values from density plot, represented by the x-axis value during highest peak, shown in red lines, were unaffected when thresholds were set higher than the minimum background intensity at 250 until 1500, where lower threshold results in appearing of false recognition from the background, indicated by the red arrows over the confocal images. The intensity for x-axis is the raw intensity obtained from images.

2.6.2 Signal intensity analysis

Signal quantifications were carried out in **Chapter 4** and **5** to determine the respective protein abundances of individual carboxysome. Therefore, after recognition of the carboxysome signals, intensity profiling using both integrated intensity and peak intensity were done. The integrated intensity corresponds to the overall volume of the dome-shaped signal coloured in green, while Peak

intensity represents the height of the dome, represent by red dashed line. In **Chapter 4** the stoichiometry of carboxysomal proteins was detected by Slimfield imaging (**section 2.5.5**). The signal intensity profiling was done using integrated intensity, while due to the advantages of peak intensity in confocal recognition from confocal imaging. Comparisons of the peak and integrated intensity profiling were done indicating well lineage correlation (R^2=0.9847), suggesting no bias generated through these two profiling methods across Slimfield and confocal imaging (**Figure 2-6B)**. Also, carboxysome signal overlays were found during imaging. Different cases of two carboxysomes nearby were shown in **Figure 2-6C**. The overlaying of carboxysome signals was take into consideration to avoid over-estimation of individual carboxysome caused by merging carboxysome signals being counted as one. The analysis indicates the superiority of peak intensity over integrated intensity by ImageJ and Fiji recognition process. Overall, the overlay of carboxysomes is only a minor event considering the general spatial distribution of carboxysome within cells (Savage et al., 2010). Therefore in large populations, the over-estimation of overlay signals would not alternate the representing intensities as kernel peak values. The over-estimation could be observed in the histogram profiles in **Chapter 4** as multi-peaks on the right tail of the major peaks.

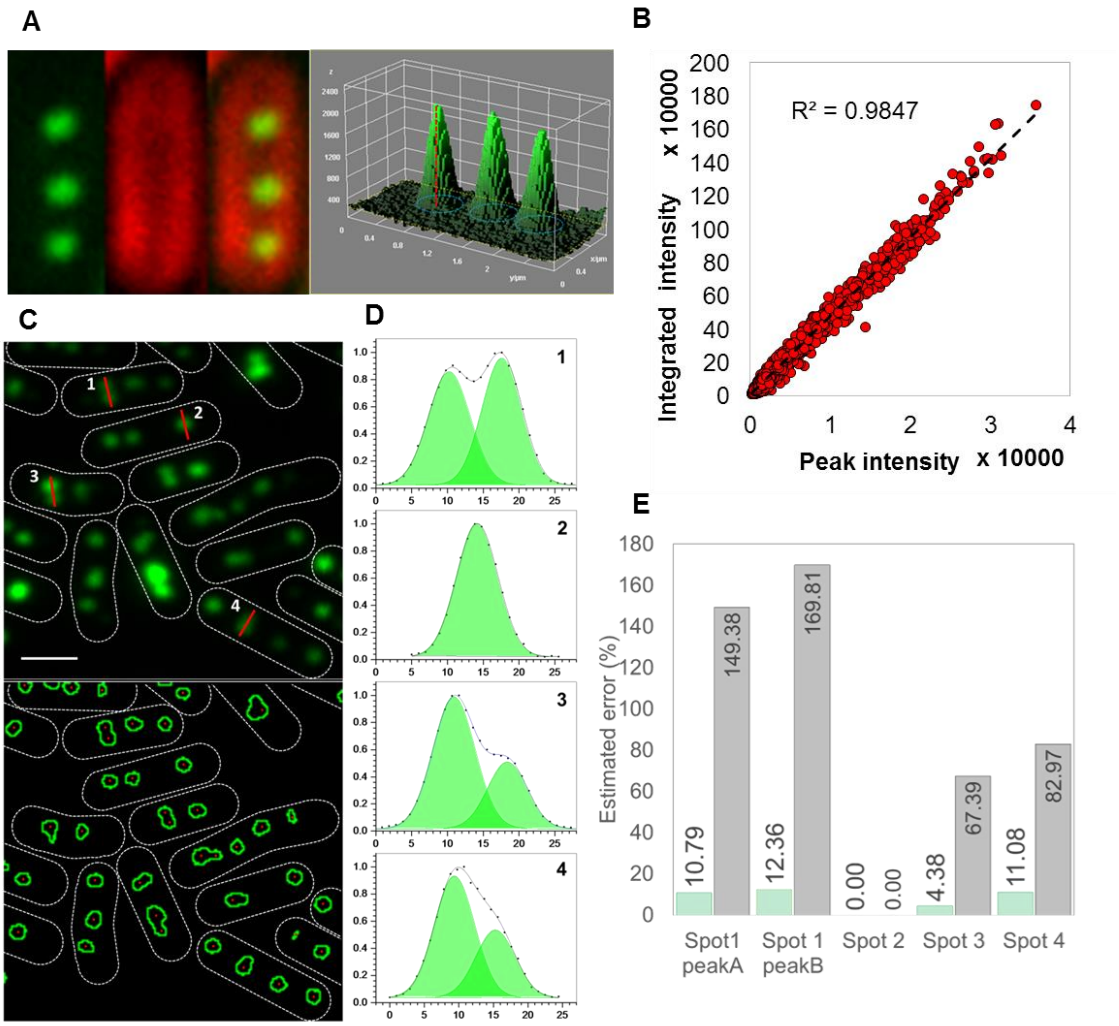


Figure 2-6. Method comparison of carboxysome signal quantification by peak and integrated intensities. A. Confocal images for RbcL-YFP under ML condition (left) and diagram for carboxysome signals (right) in which peak intensity was illustrated as red dashed line while integrated intensity represented by the volume of the green dome. **B.** 2D scatter point plot for correlated peak intensities and integrated intensities indicate lineage correlation (R^2=0.9847) of values, suggesting no bias created based the two signal types; **C.** Diagrams of recognised carboxysome through peak and integrated intensities. Top figure shows the YFP channels signals (green spot), bottom figure shows the maximum intensity (red dot) and integrated intensity border overlay (coloured as green). Four types of signal overlapping were shown as 1,2,3,4 respectively, representing Peak intensity unaffected by overlaps, no overlaps, peak intensity partially overlaps (one affected by the other) and peak intensity complete overlap (for both peaks); **D.** Profile of intensities (y-axis: relative intensities) across red line (x-axis: pixels) for 4 marked signals in A. Black line indicate the profile of signals, green coloured area are Gaussian fitted individual signal predicted from Origin; **E.** Estimate system error (percentage against theoretical values) in clustered carboxysomal signals recognised by maximum intensity (green bar) and Integrated intensity (grey bar).

Slimfield data analysis was carried out by Dr Adam Wollman. Analysis was performed using bespoke MatLab (Miller et al., 2015) with previously outlined methods (Wollman et al., 2016). In brief, candidate bright foci were identified in images using morphological transformation and thresholding. The sub-pixel centroids of these foci were determined using iterative Gaussian masking and their intensity quantified as the summed intensity inside a 5-pixel radius region of interest (ROI) corrected for the mean background intensity inside a surrounding 17x17 pixel square (ROI). Foci were accepted and tracked through time if they had a signal to noise ratio, defined as the mean intensity in the circular ROI divided by the standard deviation in the outer ROI, over 0.4. The characteristic intensity of single eYFP/mYpet was measured from the distribution of detected foci intensity towards the end of the photobleaching, confirmed by comparing the obtained value to individual photobleaching steps (**Figure 4-3**). The stoichiometry of a focus was then determined by dividing its initial intensity by the characteristic YFP intensity.

For high copy number strains, intensity of carboxysomes was very high compared to the chlorophyll but for CcmL (typically ~2x, compare **Figure 4-4** with **Figure 4-5A**) the fluorescence intensity per carboxysome was comparable (although generally brighter) to small regions of bright chlorophyll, detected as by our software, as confirmed by looking at the parental strain with no YFP present. To correct for this chlorophyll content, we tracked parental WT Syn7942 cells as YFP labelled cells to calculate the apparent chlorophyll stoichiometry distribution (**Figure 4-5A**). The CcmL distribution was then corrected by subtracting the apparent chlorophyll distribution. To investigate putative periodic features in the stoichiometry distribution, we used the raw uncorrected values to minimise dephasing artefacts (**Figure 4-5C**). Using a kernel width of 0.5 molecules (equivalent to the error in determining the characteristic intensity). The peak values in other strains were far from the chlorophyll peak and so unaffected by this correction.

2.6.3 Recognition and localisation analysis of carboxysome in Syn7942 cells

Syn7942 cells were generally recognised by chlorophyll auto-fluorescence in fluorescence microscopy. For statistical analysis for cell dimensions such as cell width and length and carboxysome numbers per cell, Image SXM software with microcompartment analysis in MIASMA package was used. The output from the software includes .csv files containing position data for carboxysome within the cell (relative value from -0.5 to 0.5) as well as cell length and width in μm (Average ± SD). Typical recognition displayed as overlay was included in the output data (**Figure 2-7)**.

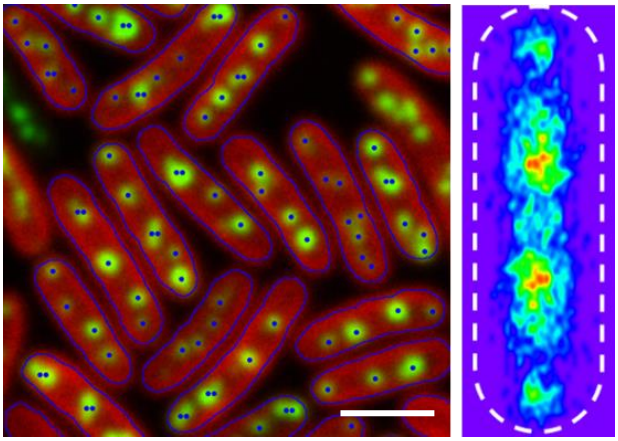


Figure 2-7. Computational programming of image analysis allows automatic identification of carboxysomes in cells in confocal images. Left, typical confocal and recognition of cell border and carboxysome (illustrated as blue lines and blue dots) by Image SXM, size bar = 2 μm; Right, the carboxysome localisation heat map within cell, where hot zone of carboxysome positions were marked out by spectra colour closer to red end while less frequent localisation was marked by spectra colour closer to blue.

The analysis for longitudinal and width localisation was done from statistics data obtained from csv files that were also provided in the software output. To further evaluate and compare profiles of distribution along the short and long axis quantitatively, two scoring parameters called spatial distribution score and polarity score were established respectively.

$$\text{The spatial distribution score} = \sum_{n=-0.5}^{0.5} (F_n - \text{Faverage})^2 * 100$$

(n is the positioning over the short axis, F_n is the frequency of positioning over position n, $F_{average}$ is the average of F_n) for which a greater value indicates stronger spatial control along the long axis of the cell.

$$\text{Polarity Score} = \sum_{i=0}^{0.5} F_i * i$$

(i is the positioning over long axis where 0 and 0.5 indicate polar and midpoint of the cells. F_i is the relative frequency of positioning over position i), where high values indicate a preference for closer to pole localisation.

2.6.4 Tracking of carboxysome in time-lapse imaging

Images were initially corrected for horizontal drifting by Descriptor-based series registration (2d/3d+T) plugin, and then were processed by the Trackmate plugin in FIJI for particle tracking. Retrieved track data was analyzed using bespoke MATLAB (Mathworks) scripts for MSD, diffusion coefficient calculations and data visualization. Diffusion coefficients were calculated by fitting to the first 6 points of the MSD *vs.* tau trace. As the MSD *vs.* tau traces indicated potentially non-Brownian diffusion over higher time interval values, we describe all diffusion coefficients as "apparent".

Chapter 3

Light modulates the biosynthesis and organisation of cyanobacterial carbon fixation machinery through photosynthetic electron flow

Chapter 3

3.1 Introduction

The efficiency of carboxysomes in enhancing carbon fixation has attracted tremendous interest in engineering the CO_2-fixing organelle in other organisms (**Chapter 1, section 1.2.3**). However, engineering of functional carboxysomes requires extensive understanding of the principles underlying the formation of β-carboxysomes and the physiological integration of β-carboxysomes into the cellular metabolism. Indeed, the carboxysome is only a part of CO_2 fixation apparatus in cyanobacterial cells. The integrated CO_2 fixation system called CCM consists of Ci uptake system (**Chapter 1, section 1.3.4**) and carboxysomes (**Chapter 1, section 1.2.1**), where substrates HCO_3^- and RuBP are continuously regenerated by Ci uptake system and CBB cycle (**Chapter 1, section 1.3.3**) in the cytoplasm.

The carboxysome-containing cyanobacteria have developed regulatory mechanism to acclimate and survive under alternated environment in nature through evolution. Particularly, light availability as a key factor to determine how much solar energy that could be harvest by photosynthetic systems (**Chapter 1, section 1.3.2**), therefore control the electron flux that drives diverse cellular metabolisms including CO_2 fixations (Masojídek et al., 2001). During a daily cycle, the light intensity varies drastically. The substrate for carboxysomes, RuBP is regenerated at different rate under different light intensities. Under irradiance at intensities below saturation, electron transport capacities are increased with higher light intensities (Tilzer, 1987), supporting faster rate of CBB cycle via the increased rate of ATP/NADPH generation (Nogales et al., 2012). Detailed regulatory mechanism of the CBB cycle is further elucidated. Key regulatory protein CP12 has been identified to regulate the CBB cycle via NADH/NADPH ratio under light/dark conditions (Tamoi et al., 2005). Therefore, enhanced RuBP regeneration could be supported by stronger electron flux under higher light intensities.

The other substrate HCO_3^- is provided by Ci transporter systems. Regarding the regulation of Ci transporters, through diversities of answering mechanism among different transporters are documented between cyanobacterial strains adapting to distinctive ecological habitats (Badger and

Price, 2003), the up-regulation of Ci transporter genes under low Ci stresses ensures the sufficient Ci intake, therefore making the CO_2 availability not to be the rate limiting factor in the CBB cycle. As light increases, Ci transporters are also upregulated to provide more HCO_3^- for carboxysomes (Gill et al., 2002; Hihara et al., 2001; Huang et al., 2002; McGinn. et al., 2004). However, when CO_2 is abundant in the environment where adequate intake is already achieved, the up-regulation by increased light is gone (McGinn et al., 2003). Besides, the large Ci influxes driven by light energy serve as electron sinks assisting in dissipation of excessive electron flow and protect cells from oxidative damage when exposed under increased light intensities (Lea-Smith et al., 2016; Xu et al., 2008). Unlike the Ci transporters, carboxysome modulation under light is poorly documented. Even though several findings have suggested the increase transcription of carboxysome genes under increased light (Gill et al., 2002; Hihara et al., 2001; Huang et al., 2002; McGinn. et al., 2004), information regarding the relationship between electron flux as direct output from light illumination and biogenesis of carboxysomes remains largely unclear.

The functional relevance of subcellular positioning of carboxysomes is another question we want to address in this chapter. Such question remains unsettled as previous works by microscopy reported distinctive localisation profiles: On one hand, evenly spaced along the centerline of the longitudinal axis of cells away from the thylakoid membrane under moderate light illumination is believed to ensure equal segregation of the machinery between daughter cells (Savage et al., 2010); meanwhile peripheral arrangement of carboxysomes near the thylakoid membranes is suggested to provide functional advantages in Ci utilisation when carboxysomes are closer to the Ci uptake system over membranes (McKay et al., 1993). Furthermore, it remains unclear how the positioning of carboxysome is regulated when carboxysomal content (in the view of this chapter, the carboxysome numbers per cell) alter under different light illuminations. It remains unclear how carboxysomes are maintained in cytoplasm when more/ fewer carboxysomes are packed within limited spaces.

In this chapter, using a combination of live-cell confocal fluorescence microscopy and biochemical and physiological approaches, we investigated the formation and spatial positioning of β-

carboxysomes in *Synechococcus elongatus* PCC7942 (hereafter Syn7942) by varying light intensities. Our study provides new insights into the regulation of β-carboxysome biosynthesis by light and the roles of photosynthetic electron flow in the carboxysome assembly. Knowledge obtained from this work is fundamental to the bioengineering and modulation of functional carboxysomes to boost photosynthetic carbon fixation in dynamic and diverse environments.

3.2 Results

We chose Syn7942 as the model organism due to its superior genetic tractability and proven suitability for fluorescence imaging (Cameron et al., 2013; Cohen et al., 2014; Liu et al., 2012; Savage et al., 2010). RbcL, the large subunit of Rubisco that resides in the β-carboxysome lumen, was tagged at the C-terminus with enhanced GFP (eGFP) and was visualised by confocal fluorescence microscopy to characterise the formation and positioning of carboxysomes in vivo. Homologous recombination was used to tag the genes at their native chromosomal locus under the control of their native promoters (**Figure 3-1**). This ensures that the fluorescently tagged proteins were expressed in context and at physiological levels.

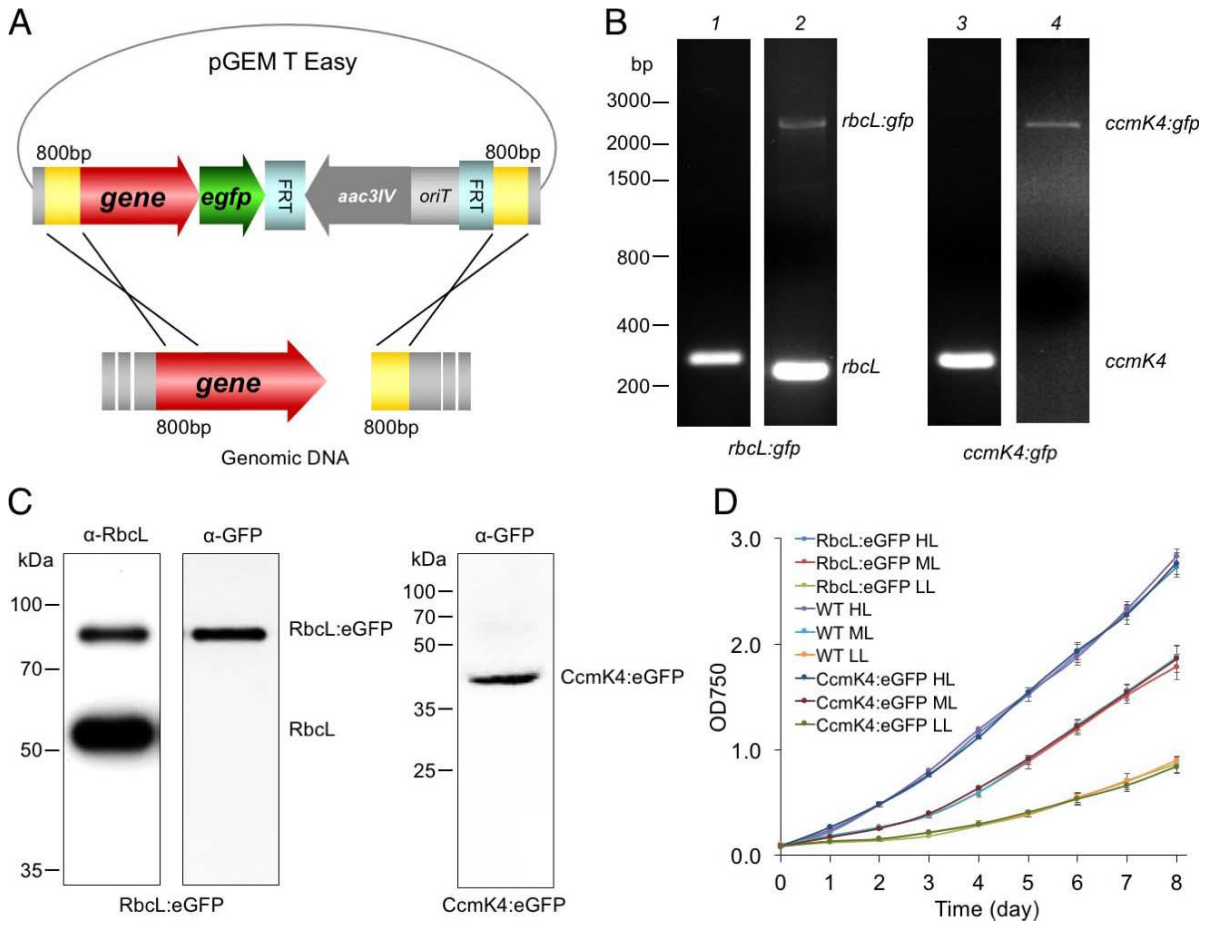


Figure 3-1. Construction and characterisation of RbcL:eGFP and CcmK4:eGFP Syn7942 strains. A, Strategy of eGFP fusion using REDIRECT protocol (described in chapter 2). **B**, PCR evaluation of RbcL:eGFP and CcmK4:eGFP genome. Lane 1 shows the *gfp* cassette fragment is fused to the C-terminus of *rbcL* in the RbcL:eGFP genome (Primers: RbcL_insert_segF, GFP 856-875 REV,). Lane 2 exhibits the *gfp* cassette fragment is partially segregated into the RbcL:eGFP genome (Primers: RbcL_insert_segF, RbcL_insert_segR). Lane 3 depicts the *gfp* cassette fragment is fused to the C-terminus of *ccmK4* in the CcmK4:eGFP genome (Primers: FccmK4_insert_segF, GFP 856-875 REV). Lane 4 illustrates the *gfp* cassette fragment is fully segregated into the CcmK4:eGFP genome (Primers: FccmK4_insert_segF, FccmK4_insert_segR). All primers were listed in **Appendix A**. **C**, Immunoblot analysis with anti-RbcL and anti-GFP antibodies of soluble fractions of the RbcL:eGFP strain and anti-GFP antibodies of soluble fractions of the CcmK4:eGFP strain, based on SDS-PAGE. Left, the band at around 83 kDa is consistent with the fusion of RbcL and eGFP, while 55 kDa refers to RbcL only. Protein quantification shows the RbcL:eGFP band is around 30% of the intensity of unfused RbcL band. Right, the 39 kDa band shows the fusion of CcmK4 and eGFP. **D**, Growth of wild-type, RbcL:eGFP and CcmK4:eGFP Syn7942 strains. The growth curves monitored at OD 750nm demonstrated no significant change of the cell growth caused by GFP fusion. Results are a mean ± SD of three independent cultures.

Figure 3-2 represents the confocal images of RbcL:eGFP Syn7942 strain. The eGFP fluorescence (green) indicates the subcellular localisation of carboxysomes, and the endogenous chlorophyll fluorescence (red) shows the organisation of thylakoid membranes. In addition, the specific DNA-staining dye 4', 6-diaminophenylindole (DAPI) was used to image chromosomes, offering the possibility to determine the cytoplasmic environment in cyanobacteria (**Figure 3-2A**). The merged channel shows that most of the cytoplasmic volume of the Syn7942 cell is densely occupied by carboxysomes and chromosomes, and no significant fluorescence gaps were visible, implying that all carboxysomes in the RbcL:eGFP transformant are likely fluorescently visible using confocal microscopy. In the free-eGFP expressing Syn7942 construct, the eGFP fluorescence is evenly spread across the cytoplasm. The distinct distributions of GFP fluorescence in the RbcL:eGFP and free-eGFP-expressing Syn7942 strains indicate the self-assembly of carboxysome proteins. PCR and immunoblot results indicate the RbcL:eGFP transformant could not be fully segregated; about 30% of total RbcL was fused with eGFP (**Figure 3-1**). The addition of GFP tag might limit the number of Rubisco proteins accommodated within the carboxysomal interior (Menon et al., 2010). Thus, there seems to be a regulation to avert full segregation and retain some unlabelled Rubisco in the carboxysome. Nevertheless, the fluorescence tagging did not affect the growth of cyanobacterial cells (**Figure 3-1**). Analysis of confocal images was programmed to examine statistically the number and spatial positioning of carboxysomes in the cell (**Figure 3-2B** and **C**, n = 300). On average, there are about four evenly positioned carboxysomes per cell, consistent with previous observations (Savage et al., 2010), confirming the physiological state of RbcL:eGFP cells. We also labelled the minor shell proteins in the carboxysome CcmK4 (Cai et al., 2015b; Kerfeld et al., 2005; Savage et al., 2010) using eGFP. PCR results demonstrate that the CcmK4:eGFP transformant was fully segregated, and the construct has similar growth rates compared to wild-type and RbcL: eGFP strains (**Figure 3-1D**).

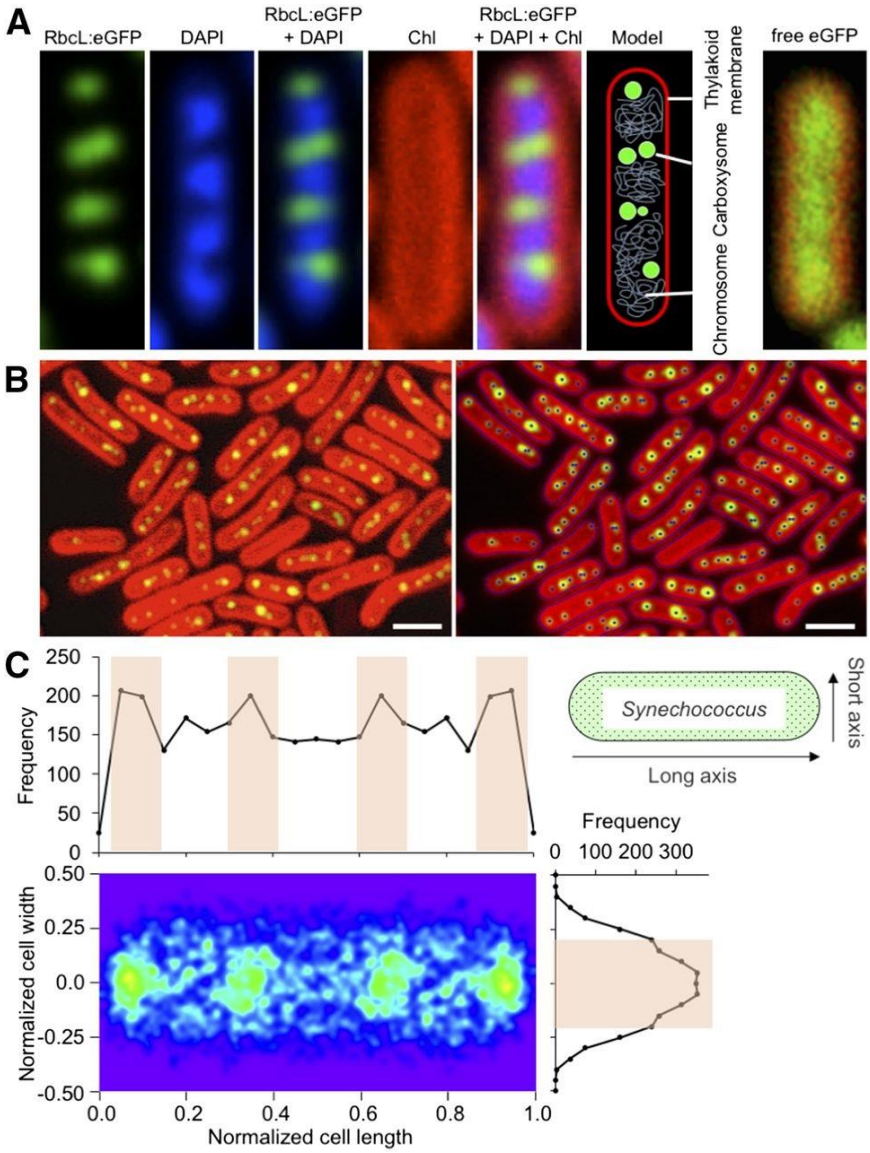


Figure 3-2. Spatial organisation of β-carboxysomes in RbcL:eGFP Syn7942 cells. A, Confocal images of a RbcL:eGFP Syn7942 cell. Green, eGFP-labelled carboxysomes; blue, DAPI-stained DNA; red, auto-fluorescence of the thylakoid membrane. The merged channel revealed that most of the cytoplasmic volume of the Syn7942 cell is occupied by carboxysomes and chromosomes. This subcellular organisation indicated that all carboxysomes in the RbcL:eGFP cell could be visualised using confocal microscopy. The confocal image of the Syn7942 construct that expresses free eGFP illustrates that free eGFP is spread throughout the cytoplasm without specific aggregation. **B**, Computational programming of image analysis allows automatic identification of carboxysomes in cells in confocal images. Bar = 2 μm. **C**, Statistical determination of the spatial localisation of carboxysomes within the cell revealed the distribution profiles of carboxysomes along both the longitudinal and short axes of the cell (n = 300). The orange squares represent the relative frequency of carboxysome localisation in the cell. The developed automated analysis software routines were used in this work for analysing the carboxysome content and positioning.

To examine whether the GFP-labelled carboxysomes can be physiologically regulated within cells, we assayed the impact of CO_2 concentration on the formation of carboxysomes. Previous studies have indicated that the carboxysomes content is affected by CO_2 availability (Harano et al., 2003; McKay et al., 1993; Whitehead et al., 2014; Woodger et al., 2003). Our confocal images show a striking reduction in the numbers of carboxysomes in cells aerated with 3% CO_2, compared to those in cells grown in ambient air (**Figure 3-3**). It reveals the feasibility of using live-cell confocal imaging to monitor the in vivo regulation of carboxysome biosynthesis in response to environmental change.

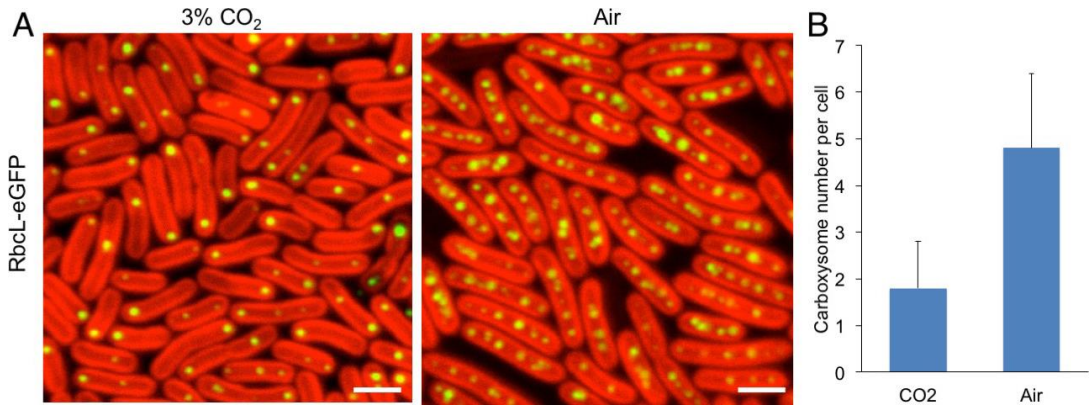


Figure 3-3. Regulation of carboxysome biosynthesis in Syn7942 by CO_2. A, Confocal images of RbcL:eGFP strain confirms that 3% CO_2 suppresses the biosynthesis of carboxysomes relative to air. Scale bar: 2 μm. **B**, Average number of carboxysomes per cell in 3% CO_2 is lower than that in air ($P < 0.05$, $n = 100$). Error bars represent SD.

3.2.1 Light triggers carboxysome biosynthesis

We studied the spatial distribution of carboxysomes in Syn7942 under the variation of light intensity: low light (LL; 10 $\mu E \cdot m^{-2}\ s^{-1}$), moderate light (ML; 50 $\mu E \cdot m^{-2}\ s^{-1}$), and higher light (HL; 100 $\mu E \cdot m^{-2}\ s^{-1}$). Confocal images of RbcL:eGFP cells show that the carboxysome abundance per cell has a strong correlation with the illumination intensity during cell growth (**Figure 3-4A**). The number of carboxysomes per cell is higher under HL, whereas LL leads to the reduction in carboxysome numbers. The light dependence of carboxysome content was further substantiated by transmission electron microscopy results of wild-type Syn7942 cells (**Figure 3-4B**; **Figure 3-5**). The numbers and positioning of carboxysomes in the cell were statistically analysed based on the confocal images.

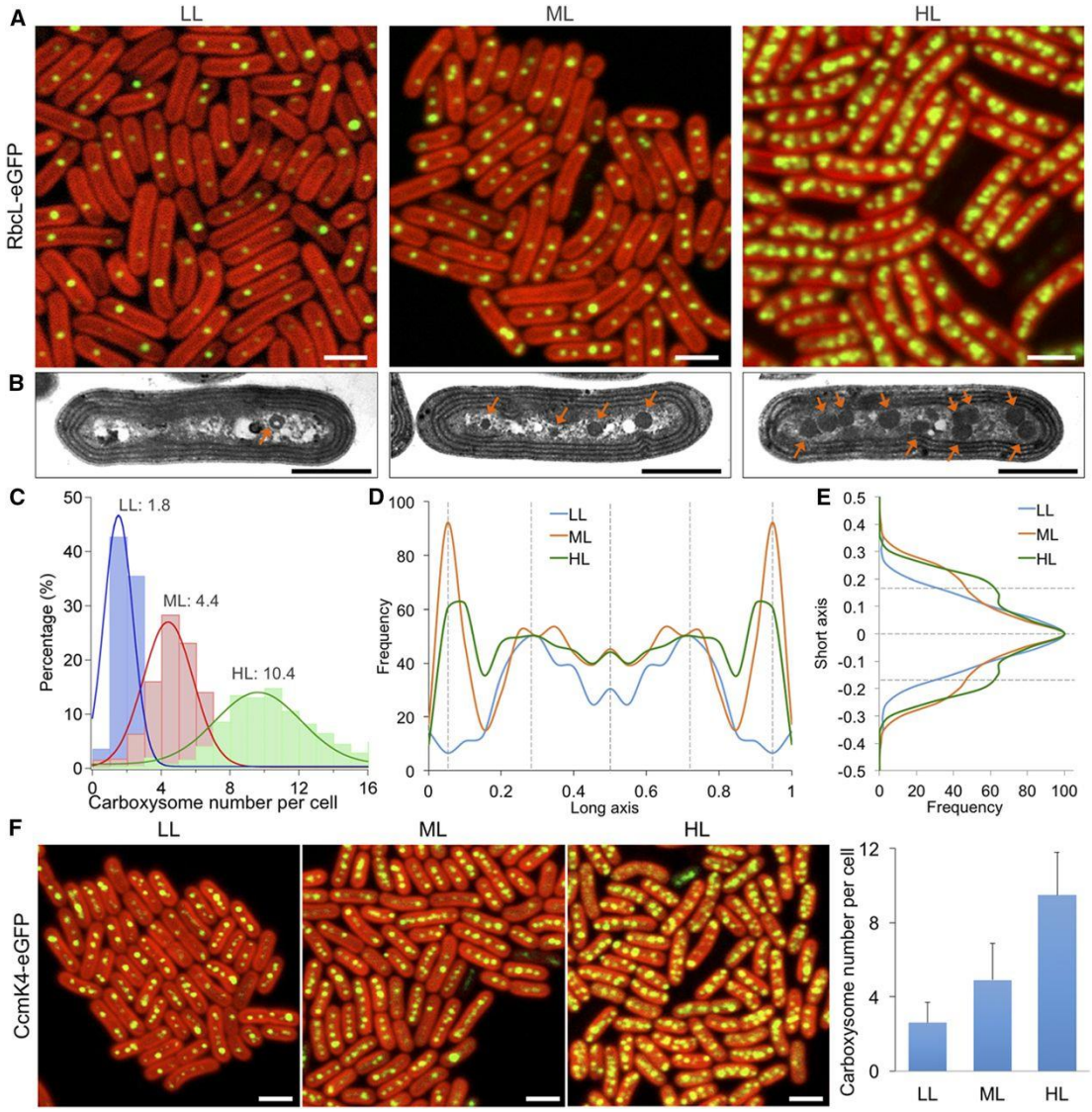


Figure 3-4. Distinct distribution patterns of carboxysomes in Syn7942 under different light intensities. A, Confocal microscopy images of the spatial organisations of carboxysomes in the RbcL:eGFP strain grown under LL (10 μE·m^{-2} s^{-1}), ML (50 μE·m^{-2} s^{-1}), and HL (100 μE·m^{-2} s^{-1}). Variations in the carboxysome content were observed under different light intensities. Bar = 2 μm. **B**, Thin-section transmission electron microscopy images of wild-type Syn7942 cells grown under LL, ML, and HL (**Figure 3-5**). The thylakoid membranes form regular multiple parallel layers surrounding the cytoplasm. The carboxysomes were observed as dark polyhedral particles (arrows) in the cytoplasm. Stronger light leads to the increase in carboxysome numbers in cells. Bar = 1 μm. **C**, Computational analysis of confocal images shows the average numbers of carboxysomes per cell under LL, ML, and HL (n = 500 for each condition). **D**, The positioning of carboxysomes along the normalised longitudinal axis of the Syn7942 cell under LL, ML, and HL. The relatively periodic and polar localisation of carboxysomes in cells are indicated (n = 500). **E**, The positioning of carboxysomes along the normalised short axis of the Syn7942 cell (n = 500). HL gives rise to a wider distribution of carboxysomes from the centerline of cells. **F**, Confocal microscopy images of CcmK4:eGFP cells show the spatial positioning of carboxysomes grown under LL, ML, and HL. Variations in the carboxysome content were observed under different light intensities, in good agreement with those of the RbcL:eGFP strain. Error bars represent SD (n = 500 for each condition). Bar = 2 μm.

On average, around two carboxysomes per cell (1.8 ± 1.2, *n* = 500) were observed under LL, whereas about four carboxysomes per cell (4.4 ± 1.9, *n* = 500) under ML and over ten (10.4 ± 3.8, *n* = 500) under HL were detected (**Figure 3-4C**). The data are in good agreement with the results from electron microscopy images (**Figure 3-5**): 1.6 ± 0.7 (LL, *n* = 30), 3.9 ± 0.8 (ML, *n* = 30), 10.2 ± 2.0 (HL, *n* = 30). No significant changes in cell dimensions were detected under different light intensities (**Figure 3-6**).

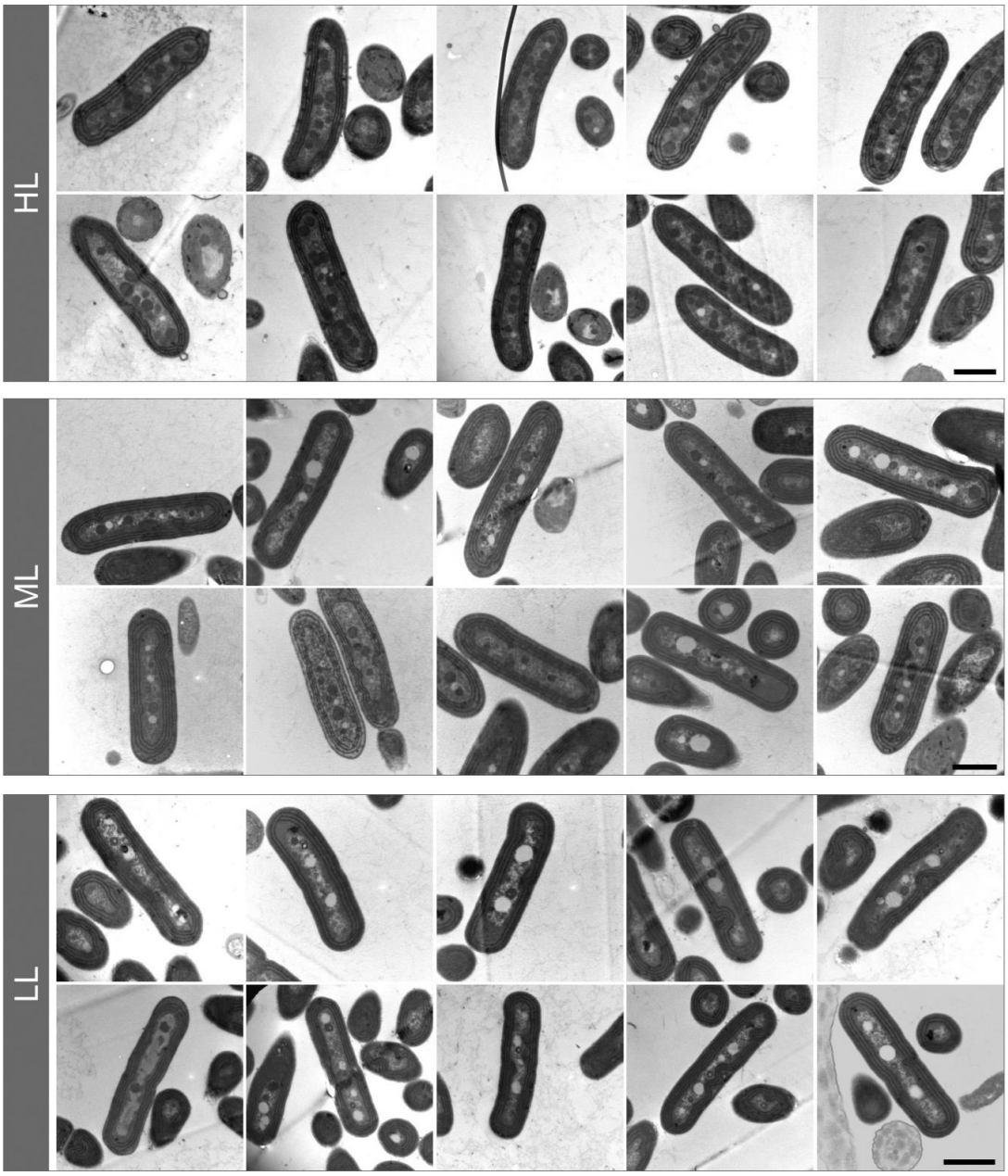


Figure 3-5. Thin-section transmission electron microscopy images of wild-type Syn7942 cells grown under HL, ML and LL. The average numbers of carboxysomes per cell are 10.2 ± 2.0, 3.9 ± 0.8 and 1.6 ± 0.7, respectively (± SD, *n* = 30 for each light condition). Scale bar: 1 μm.

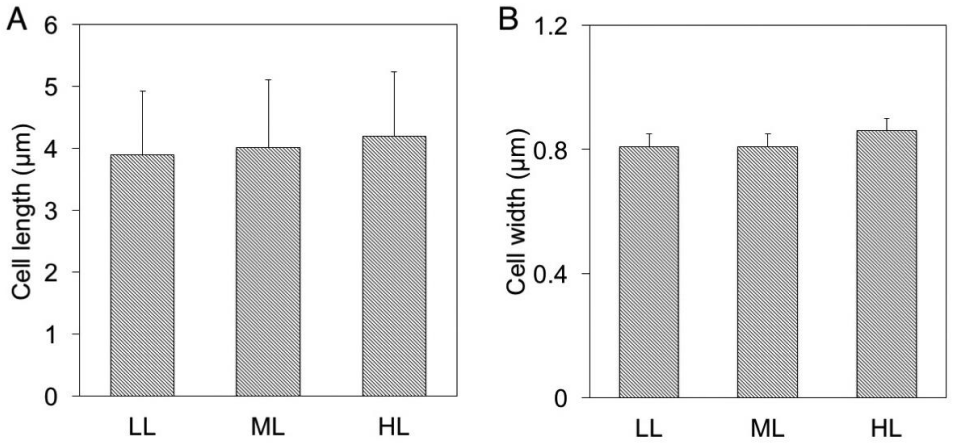


Figure 3-6. The sizes of Syn7942 cells remain similar under the variation of light intensity. A, Averaged cell lengths are similar in LL, ML and HL ($P > 0.05$, $n = 500$). **B**, Difference in the averaged cell widths is not detectable in LL, ML and HL ($P > 0.05$, $n = 500$). Error bars represent SD.

These results indicate that light intensity plays an important role in determining the biosynthesis of β-carboxysomes in Syn7942. Varying light intensities could also result in different organisational patterns of carboxysomes in cells. Image analysis reveals even distribution of carboxysomes along the longitudinal axis of the cell (**Figure 3-4D**). Carboxysomes tend to locate at approximately one-fourth position along the cell length under LL, whereas under ML and HL a polar location of carboxysomes within the cell was observed apart from the even distribution (**Figure 3-4D**). Analysis of the positioning of carboxysomes along the cell width elucidates that increasing numbers of carboxysomes induced by stronger light present a wider distribution along the short axis of the cell, compared to the centre line positioning observed under LL (**Figure 3-4E**). This organisation likely provides a means to house more carboxysomes in a spatially crowded cytoplasm environment. Consistent with the results of the RbcL:eGFP construct, our confocal images of the CcmK4:eGFP construct also show the increase in carboxysome content triggered by stronger irradiance (**Figure 3-4F**). Similarly, light regulation of carboxysome content in Syn7942 was also seen in the RbcL:YFP strain (Savage et al., 2010) in which rbcL:yfp was inserted into a neutral site rather than the native locus in the genome (**Figure 3-7**).

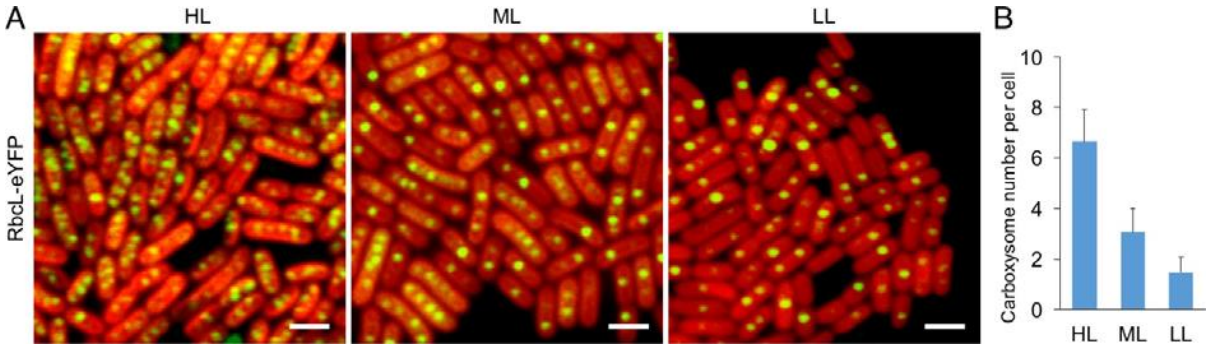


Figure 3-7. Light regulation of carboxysome content in RbcL:YFP cells. A, Confocal microscopy images of the RbcL:YFP Syn7942 construct (pDFS621, PapcA:rbcL:YFP, Savage et al., 2010) grown at HL, ML and LL. **B**, the carboxysome content within the Syn7942 cells is dependent on light intensity, in agreement with the observation of RbcL:eGFP and CcmK4:eGFP cells. Error bars represent SD (n = 250). Scale bar: 2 μm.

Together, our observations reveal a general regulation of carboxysome content and organisation in Syn7942 in response to variations in light intensity. The light-regulated carboxysome biosynthesis was further characterised by time-lapse confocal imaging during cell growth (**Figure 3-8**). HL treatment on cells that were preadapted to LL resulted in a linear increase in carboxysome content over five days. Reversibly, LL treatment caused a reduction in carboxysome numbers, although the rate of reduction is lower than that of the increase in carboxysome numbers. These results indicate that the light-dependent carboxysome biogenesis might function as a long-term acclimation process in cyanobacteria. On the other hand, despite potential repairing mechanism led by protein dynamics (Sutter et al., 2016), there appears to be no specific degradation pathway for carboxysomes. The stability of mature carboxysomes in vivo may be of physiological importance for the cellular metabolism (Cameron et al., 2013).

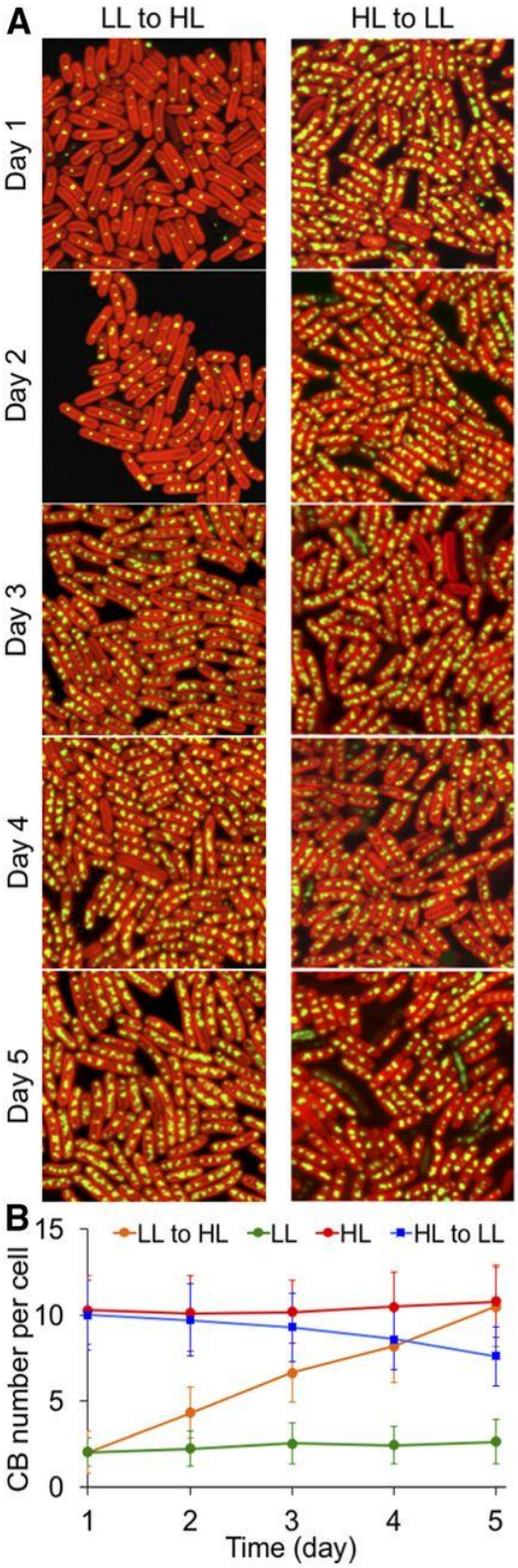


Figure 3-8. Characterisation of the light-regulated biosynthesis process of β-carboxysomes. A, Time-lapse confocal images of the LL-adapted RbcL:eGFP Syn7942 strain under HL treatment and the HL-adapted RbcL:eGFP strain under LL treatment. Cells from the same flasks were imaged under confocal microscopy once per day, for 5 days continuously. Changes in carboxysome content per cell were captured. **B**, Analysis of the average numbers of carboxysomes based on confocal images reveals an increase in carboxysome abundance induced by HL and a decline in carboxysome abundance caused by LL, compared to the numbers of carboxysomes under constant LL or HL treatments. Error bars represent SD (n = 250).

3.2.2 Light-induced carboxysome biosynthesis determines the carbon fixation activity of cells

In addition to the confocal microscopy results that reveal the light-induced carboxysome content in the cell, immunoblot analysis show that the abundance of Rubisco proteins per cell, normalised using the AtpB content (Zhang et al., 2012), is also up-regulated by increasing irradiance (**Figure 3-9A**). The Rubisco abundance under HL is about 6 and two times as high as those under LL and ML, respectively (**Figure 3-9B and C**).

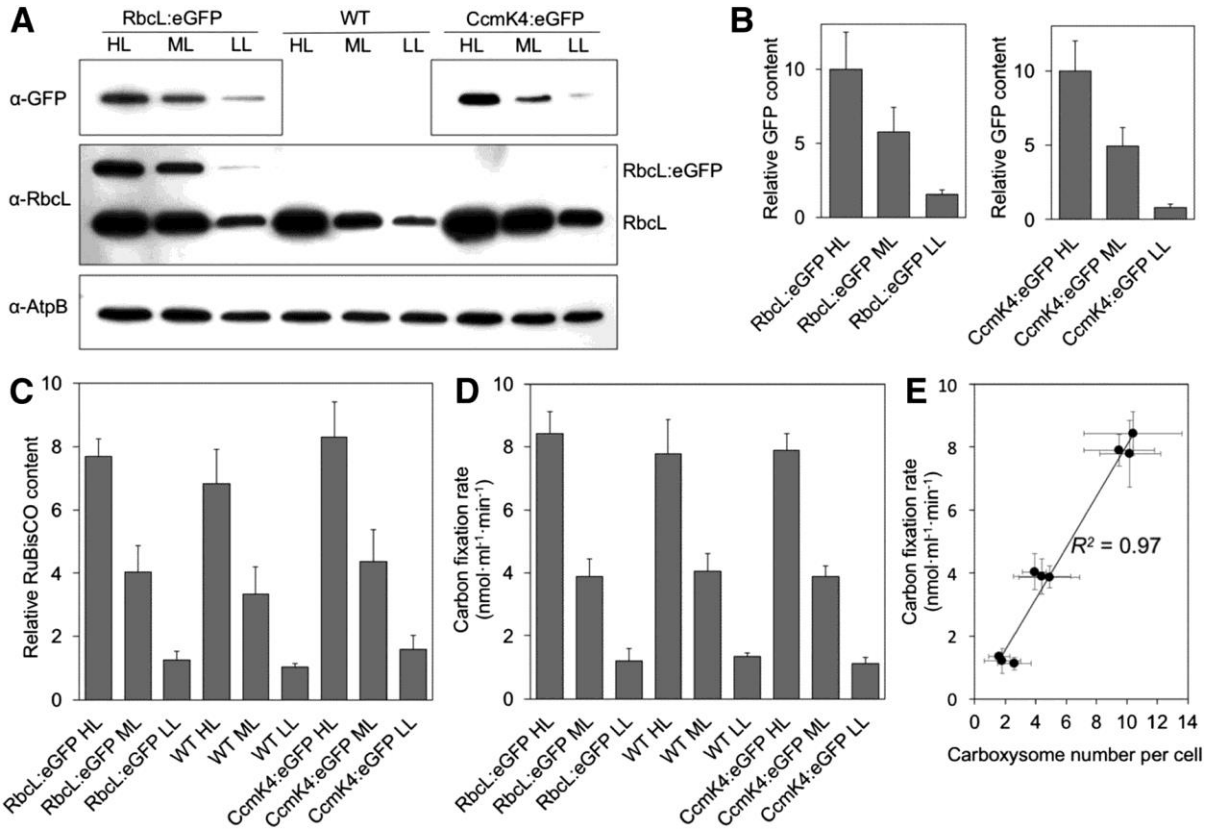


Figure 3-9. Light modulates the biosynthesis of β-carboxysomes and cellular carbon fixation. A, Immunoblot analysis using anti-RbcL and anti-GFP antibodies shows variations of the Rubisco content in RbcL:eGFP cells. HL triggers the accumulation of Rubisco. Using anti-RbcL antibody, RbcL:eGFP strains present two bands: the upper band for RbcL-eGFP and the lower band for RbcL only. AtpB was used as a loading control (Zhang et al., 2012). Gels are representative of six independent experiments. **B**, Immunoblot analysis implies that the GFP amount in RbcL:eGFP and CcmK4:eGFP cells varies under different light intensities (±SD, $n = 6$, $P < 0.05$). **C**, Densitometry of RbcL in wild-type, RbcL:eGFP, and CcmK4:eGFP cells is dependent on light intensity (±SD, $n = 6$, P

< 0.05). **D**, 14C carbon fixation rates of wild-type, RbcL:eGFP, and CcmK4:eGFP cells under LL, ML, and HL at 0.5 mM RuBP (±SD, n = 6). The cell density was normalised using the AtpB content (**Figure 3-9A**). The carbon fixation rates of cells as a function of RuBP are shown in **Figure 3-11**. **E**, The carbon fixation rate per cell is proportional to the numbers of carboxysomes within the cell (R^2 = 0.97). The numbers of carboxysomes per cell were determined from electron microscopy images for the wild-type strain and confocal microscopy images for RbcL:eGFP and CcmK4:eGFP strains.

Interestingly, a slight increase in Rubisco content was observed in both RbcL:eGFP and CcmK4:eGFP strains compared to wild-type cells, probably as compensation for compromised CO_2-fixing activities of carboxysomes caused by fluorescence labelling. The light-induced changes in Rubisco content at the protein level was also confirmed by measuring the total fluorescence intensity of RbcL:eGFP per cell (**Figure 3-10**).

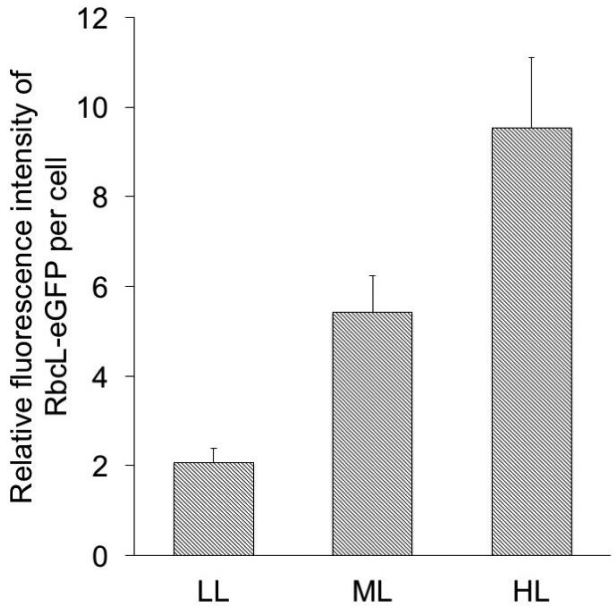


Figure 3-10. Relative abundance of Rubisco in RbcL:eGFP Syn7942 strain under LL, ML and HL, based on confocal image analysis. Analysis of the relative fluorescence intensity of RbcL:eGFP per cell based on confocal images (**Figure 3-4**) exhibits the increase of GFP intensity in the Syn7942 cell with the rise of illumination intensity, indicating that stronger illumination could stimulate the biosynthesis of Rubisco enzymes (n = 100). It is further confirmed by immunoblot analysis shown in **Figure 3-9**. Error bars represent SD.

Our results corroborate previous studies, which revealed that the transcription of carboxysome genes is stimulated in response to increasing light intensity (Gill et al., 2002; Hihara et al., 2001; Huang et al., 2002; McGinn. et al., 2004). To verify the physiological coordination between carboxysome content and carbon fixation in Syn7942, we surveyed the carbon fixation activities of cells (based on the AtpB content) under different light conditions. To examine the maximum carbon fixation rates, 0.5 mM D-ribulose 1, 5-bisphosphate sodium salt hydrate (RuBP) was applied (**Figure 3-11**).

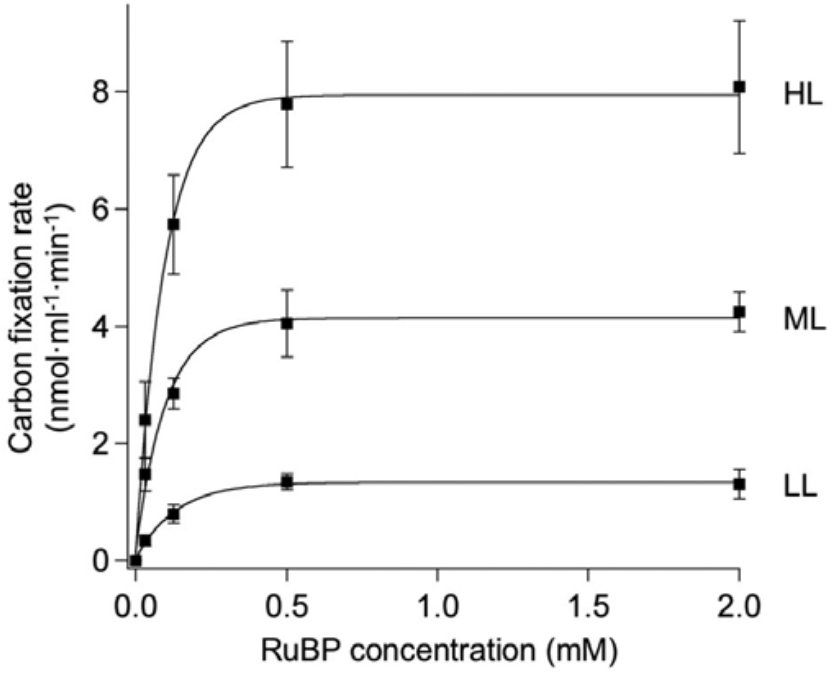


Figure 3-11. ^{14}C carbon fixation rates of wild-type Syn7942 cells grown under LL, ML and HL, as a function of RuBP concentration (± SD, *n* = 6). The carbon fixation rate of wild-type cells in HL (7.8 ± 1.1 nmol min^{-1} ml^{-1}) is higher compared with those in ML (4.0 ± 0.4 nmol min^{-1} ml^{-1}) and LL (1.3 ± 0.3 nmol min^{-1} ml^{-1}). The curves were fitted exponentially. The cell density was calibrated using the AtpB content of cells. 0.5 mM RuBP was used during Rubisco assay in this study to determine the maximum carbon fixation rates (see **Figure 3-9**).

Figure 3-9 depicts a strong dependence of carbon fixation rates of Syn7942 cells on light intensity. The carbon fixation rate of wild-type cells in HL (7.8 ± 1.1 nmol min^{-1} ml^{-1}) is higher compared with those in ML (4.0 ± 0.4 nmol min^{-1} ml^{-1}) and LL (1.3 ± 0.3 nmol min^{-1} ml^{-1}) shown in **Figure 3-9D**). Similar tendency was also observed in RbcL:eGFP and CcmK4:eGFP cells, indicating that increasing irradiance enhances the carbon fixation of Syn7942 cells (**Figure 3-9D**). Furthermore, there is a close

correlation between the numbers of carboxysomes and carbon fixation rate of cells (**Figure 3-9E**). Together, our results indicate explicitly that the light-intensity-regulated carboxysome biosynthesis serves as a regulatory mechanism of modulating the capacity of CO_2 fixation in the cell.

3.2.3 Light regulation of carboxysome biosynthesis is mediated by photosynthetic electron flow

Changes in light intensity could alter electron flow and redox states of intersystem electron carriers, especially the Plastoquinone (PQ) pool (Liu et al., 2012; Mullineaux, 2001). We conducted extensive studies on the carboxysome formation process in response to irradiance variations in the presence of two specific inhibitors of photosynthetic electron transport, namely, 3-(3,4-dichlorophenyl)-1,1-dimethylurea (DCMU) and 2,5- dibromo-3-methyl-6-isopropyl-benzoquinone (DBMIB). DCMU and DBMIB inhibit photosynthetic electron transport from the photosystem II complex to the PQ pool and from the PQ pool to the cytochrome b6f complex, respectively (Trebst, 1980). We observed that when light is switched from LL to HL for 24 h, both DCMU and DBMIB treatments suppress the O_2 evolution from photosystem II of cells (**Figure 3-12**) and hampers the light-induced carboxysome biosynthesis (**Figure 3-13**).

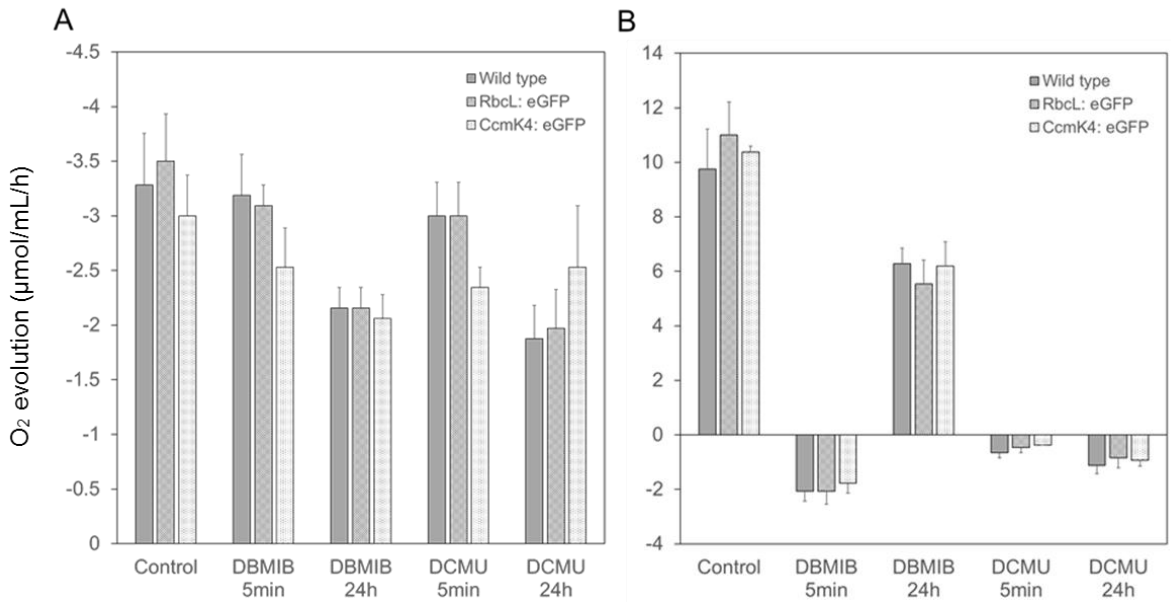


Figure 3-12. Oxygen consumption and evolution with inhibitor DBMIB and DCMU treatment at five mins and 24 hours of treatment during dark (A) and light (B) conditions with HL-adapted cells. Error bars represent SD (n = 4). Oxygen evolution analysis of Syn7942 cells in the presence of DBMIB and DCMU for 24 hours. Error bars represent SD (n = 4). The oxygen evolution

was inhibited in the presence of DBMIB for 24 hours and was not detectable in the presence of DCMU.

It demonstrates that the inhibition of photosynthetic electron flow impairs the biogenesis of carboxysomes in cyanobacteria. The increase in carboxysome content in the cell was also inhibited by lincomycin (**Figure 3-13**), a protein synthesis inhibitor that suppresses de novo protein synthesis (Dalla Chiesa et al., 1997). The similar effects of DCMU, DBMIB, and lincomycin on impeding the carboxysome biosynthesis suggest that the regulation of photosynthetic electron flow may affect the synthesis and assembly of carboxysome proteins to form carboxysomes.

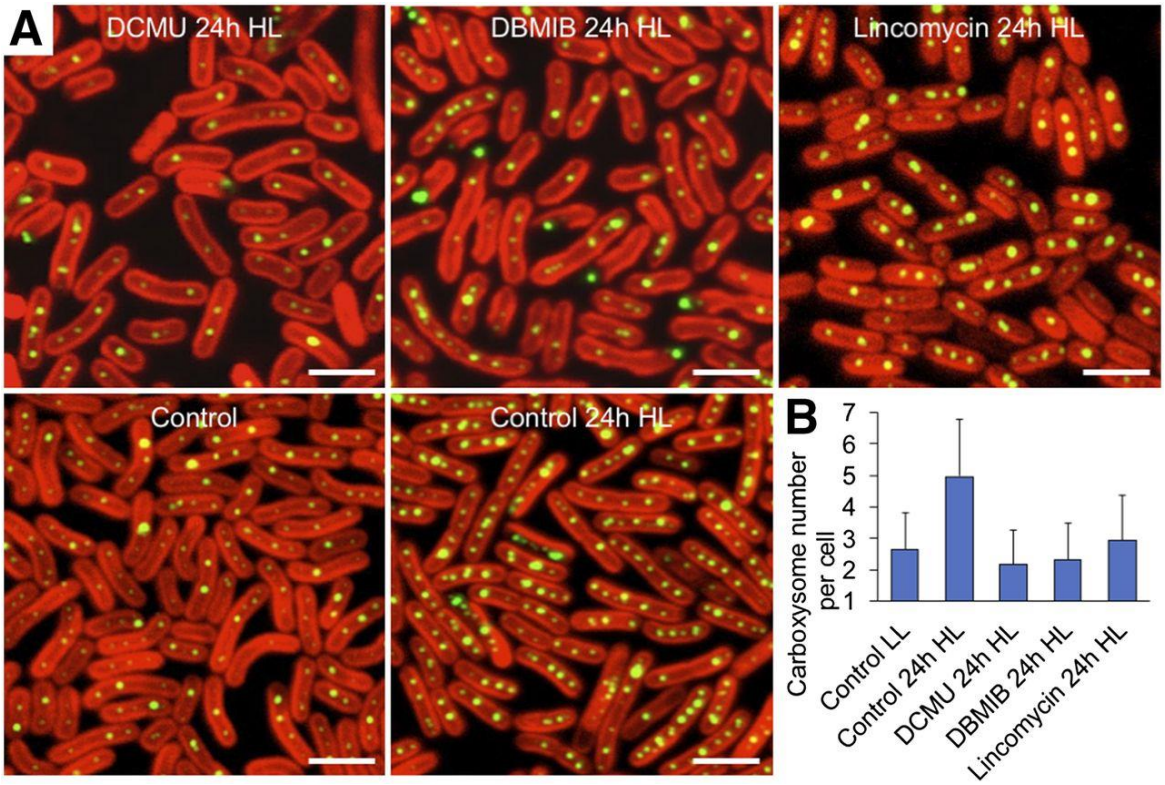


Figure 3-13. β-carboxysome biosynthesis is regulated by photosynthetic electron flow. A, Confocal microscopy images of LL-adapted RbcL:eGFP Syn7942 cells after 24-h HL treatment, in the presence of the photosynthetic electron transport inhibitors DCMU and DBMIB or the protein synthesis inhibitor lincomycin. The control images were captured in LL-adapted Syn7942 cells grown at LL and after 24-h HL treatment without inhibitor treatments. Bar = 2 μm. **B**, Analysis of the average numbers of carboxysomes per cell (±SD, *n* = 300) based on the confocal images illustrates the suppression of carboxysome content with the treatments of DCMU and DBMIB,

indicating that light intensity regulates the biosynthesis and assembly of carboxysomes through photosynthetic electron flux.

3.2.4 Carboxysome localisation is sensitive to the redox state of photosynthetic electron transport

Closer inspection of the distribution of carboxysomes in HL-adapted cells illustrates that along with the dense packing of carboxysomes in the cytoplasm, a few carboxysomes aggregate preferentially and form a single large carboxysome "cluster". Several clusters are then evenly positioned along the longitudinal axis of the cell (**Figure 3-14**). It was postulated that the local CO_2 concentration near each carboxysome is higher (Mangan and Brenner, 2014). The carboxysome clusters may be functionally advantageous to minimising CO_2 leakage and maximising the CO_2 accumulation around all carboxysomes in the cytoplasm (Ting et al., 2007), thereby enhancing carbon fixation of cells. The equally spaced carboxysome clusters along the long axis of the cell may correlate with the positioning of chromosomes (Jain et al., 2012). It could favour the equal segregation of carboxysomes between daughter cells during cell division (Savage et al., 2010). We further surveyed the effects of DCMU and DBMIB on the spatial organisation of carboxysomes in cells. In HL-adapted RbcL:eGFP cells, carboxysomes possess a dense distribution, suitable for detecting the spatial redistribution of carboxysomes.

Under DCMU treatment, carboxysomes present the typical "clustering" distribution in the cytoplasm (**Figure 3-14B**), similar to the observations in non-treated cells. By contrast, DBMIB treatment gave rise to a linear positioning of carboxysomes along the centerline of the cell (**Figure 3-14C**). Image analysis reveals that, relative to the marked repositioning of carboxysomes, no detectable differences in the thylakoid membrane structure and cytoplasmic volume (indicated by endogenous chlorophyll fluorescence) were observed (**Figure 3-14D**), and whereas the carboxysome numbers per cell under DCMU and DBMIB treatments were comparable (**Figure 3-14E**). The periodic distribution of carboxysomes along the long axis of DBMIB-treated cells is somewhat less significant than that of DCMU-treated cells, and the polar localisation of carboxysomes seems to disappear in DBMIB-

treated cells (**Figure 3-14F** and **G**). More remarkable changes were observed in carboxysome distribution along the short axis of the cell (**Figure 3-14F** and **H**). DBMIB treatment led to the relocation of carboxysomes from a wider distribution along the cell width to a narrow positioning at the centerline of the short axis of the cell, compared to the wide distribution observed in DCMU-treated and non-treated cells (**Figure 3-14H**). DCMU and DBMIB have opposite effects on the redox state of the PQ pool in photosynthetic electron transport chain: The PQ pool is oxidised by DCMU and reduced by DBMIB. The distinct effects of DCMU and DBMIB indicate that the spatial organisation of β-carboxysomes in Syn7942 correlates with the redox state of photosynthetic electron transport chain. Consistent with the finding of HL-adapted cells, the reorganisation of carboxysomes under DCMU and DBMIB treatments was also observed in LL- and ML-adapted RbcL:eGFP cells (**Figure 3-15**) and CcmK4:eGFP cells (**Figure 3-16**), corroborating that the redox regulation of photosynthetic electron flow could affect the spatial positioning of β-carboxysomes in Syn7942.

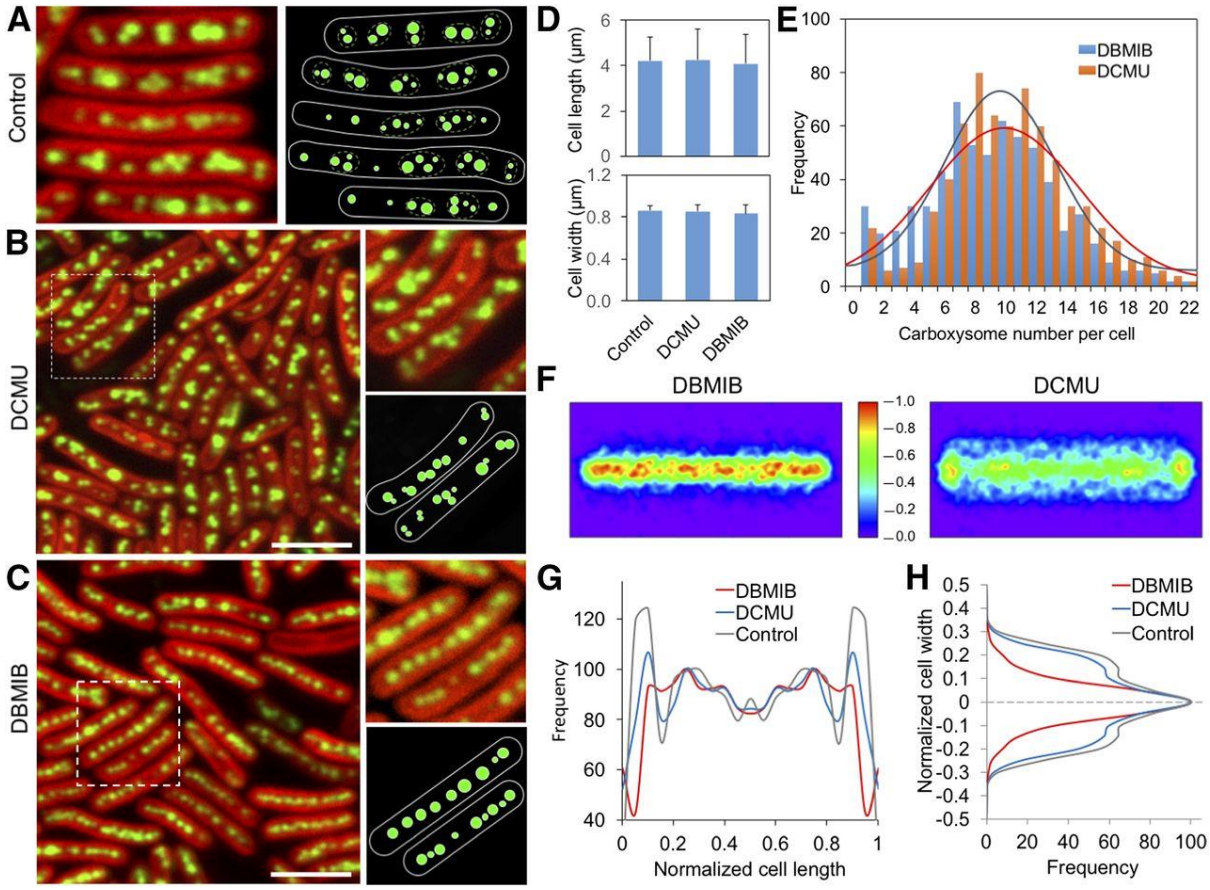


Figure 3-14. Correlation between the spatial organisation of β-carboxysomes and redox state of photosynthetic electron transport chain in Syn7942. A, Confocal image (left) and schematic model (right) of the HL-adapted RbcL:eGFP strain shows the evenly distributed carboxysome clusters (orange dotted circles), each of which contains several carboxysomes assembled together in the local cytoplasmic region. **B**, Confocal image of HL-adapted RbcL:eGFP cells under DCMU treatment shows the clustering distribution of carboxysomes along the centre line of the cell. Bar = 5 μm. **C**, Confocal image of HL-adapted Syn7942 cells under DBMIB treatment shows the linear distribution of carboxysomes along the centre line of the cell. Bar = 5 μm. **D**, No significant changes in the cell length and width of Syn7942 is detected during inhibitor treatments for 24 h ($P > 0.05$, $n = 500$). Error bars represent sd. **E**, Average numbers of carboxysomes per cell under the treatments of DCMU and DBMIB do not have remarkable changes ($P > 0.05$, $n = 500$). **F**, Normalised spatial distribution maps of carboxysomes in cells under the treatments of DCMU and DBMIB. The bar presents the relative frequency of carboxysome localisation. **G**, The distributions of carboxysomes along the cell length under DCMU and DBMIB treatments ($n = 500$). **H**, DBMIB treatment results in a linear positioning of carboxysomes at the centerline of HL-adapted Syn7942 cells, whereas the DCMU treatment leads to a wider distribution of carboxysomes along the cell width ($n = 500$). Similar results were also obtained in LL- and ML-adapted Syn7942 cells (**Figure 3-15**) and CcmK4:eGFP cells (**Figure 3-16**).

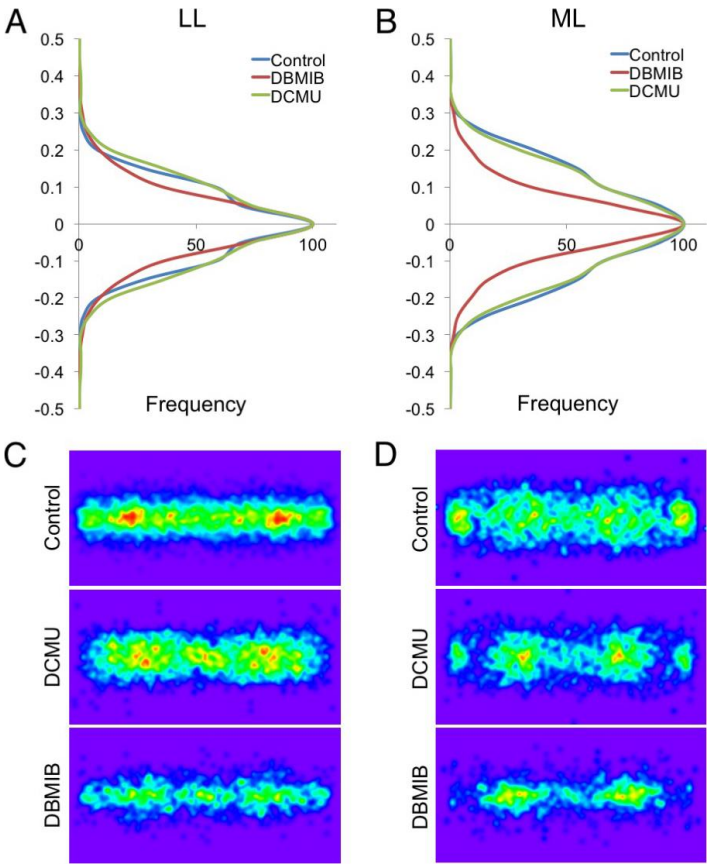


Figure 3-15. Organisation of carboxysomes in LL- and ML-adapted cells under the treatment of DCMU and DBMIB. **A** and **B**, The narrow distributions of carboxysomes along the short axis of LL- and ML-adapted cells under DBMIB treatment (*n* = 300), compared to the wider distribution induced by DCMU treatment, indicated the redox state of PQ pool plays an important role in regulating the carboxysome localisation in cells (**Figure 3-14**). **C** and **D**, Normalised spatial distribution maps of carboxysomes within LL- and ML-adapted cells under the treatment of DCMU and DBMIB.

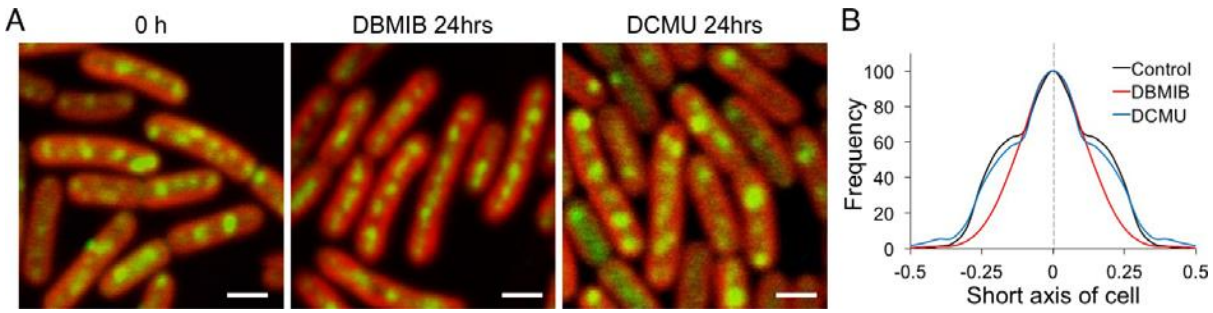


Figure 3-16. **The effects of DBMIB and DCMU on β-carboxysome localisation in CcmK4:eGFP cells. A**, Confocal images of HL-adapted CcmK4:eGFP cells under DBMIB and DCMU treatments for 24 hours. Scale bar: 1 μm. **B**, Analysis of carboxysome distribution along the short axis of cells (*n* = 200). It shows different effects of DBMIB and DCMU on the carboxysome positioning in Syn7942, consistent with the observations of RbcL:eGFP cells.

We further observed that changes in carboxysome localisation are not clearly visible within 4-h DBMIB treatment (**Figure 3-17**), implying that the reorganisation of carboxysomes seems to be a long-term adaptive process in response to redox regulation.

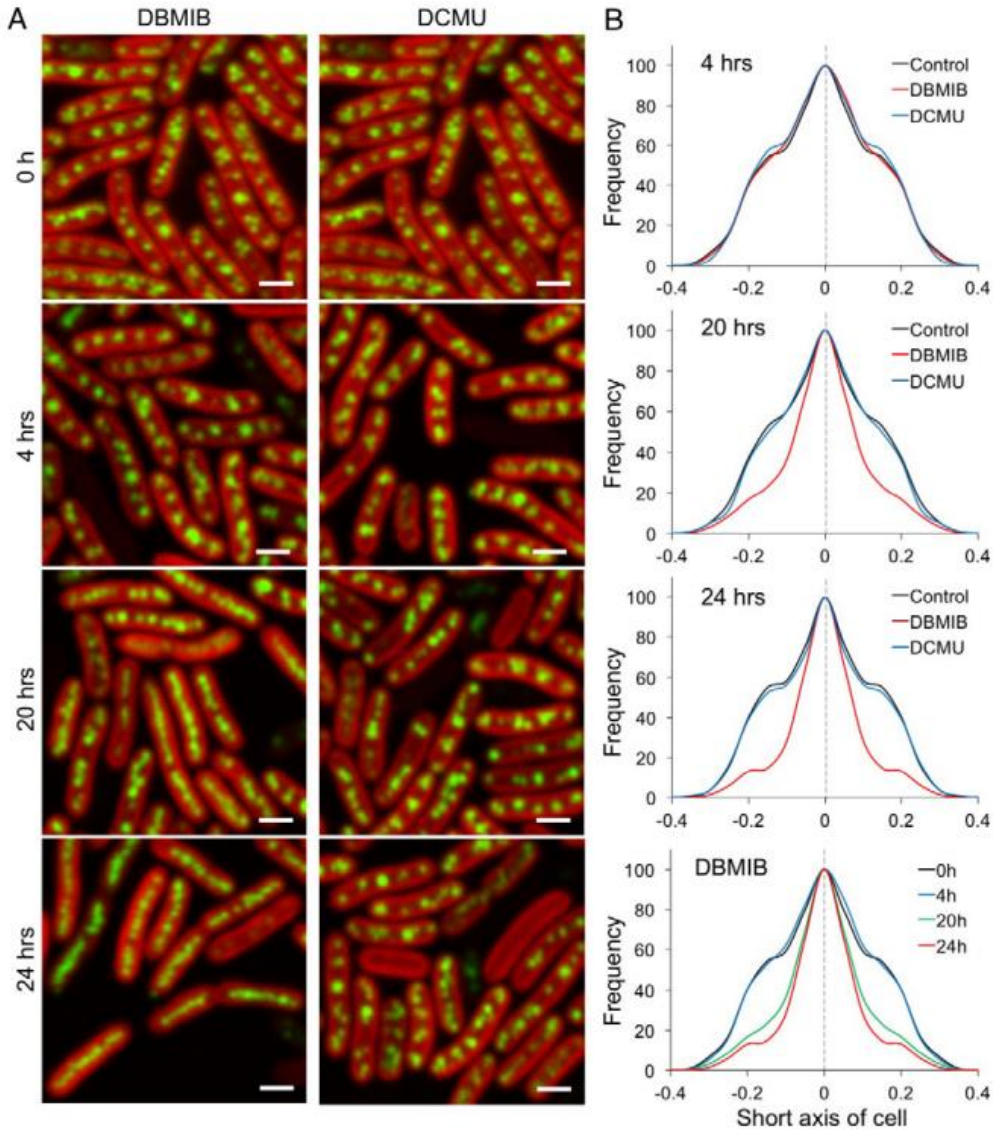


Figure 3-17. Time-lapse confocal fluorescence imaging of RbcL:eGFP cells in the presence of DCMU and DBMIB. A, Confocal images of HL-adapted RbcL:eGFP cells with DBMIB and DCMU treatments for 4, 20 and 24 hours. Scale bar: 1 μm. **B**, Carboxysome localisation along the short axis of Syn7942 cells, demonstrating different effects of DCMU and DBMIB on the spatial positioning of carboxysomes. 24-hour inhibitor treatments were applied to characterise the changes in the carboxysome distribution (**Figure 3-14**).

3.3 Discussion

In cyanobacteria, light is fundamental to energy production, DNA replication, and the regulation of gene expression (Asayama, 2006; Ohbayashi et al., 2013). Light-dependent reactions of photosynthesis generate chemical energy, in the forms of ATP and NADPH, which is utilised to drive the Calvin-Benson-Bassham Cycle responsible for CO_2 fixation into metabolisable sugars. It was found that HL could induce an increase in the transcription of Rubisco and carboxysome *ccm* genes (Gill et al., 2002; Hihara et al., 2001; Huang et al., 2002; McGinn. et al., 2004). In this work, we evaluated the impact of light intensity on the regulation of β-carboxysome biosynthesis at the whole organelle and cellular levels. Our results show that increasing irradiance triggers the accumulation of carboxysome proteins (**Figure 3-9A**; **Figure 3-10**) and formation of functional carboxysomes (**Figure 3-9**), thereby enhancing carbon fixation of cells (**Figure 3-9**). We further revealed the close correlation between light-regulated photosynthetic electron flow and β-carboxysome biosynthesis (**Figure 3-13**). Given that the expression of carboxysome genes and their encoded proteins are under light-dark regulation (Aryal et al., 2011; Ito et al., 2009; Watson and Tabita, 1996), further work needs to be directed to elucidate whether *de novo* assembly of carboxysomes is regulated and/or potentially gated by the cyanobacterial circadian rhythm, as shown in **Chapter 5**.

The spatial distribution of β-carboxysomes along the longitudinal axis of Syn7942 cells is driven by interactions with the cytoskeleton, ensuring the equal segregation of carbon fixation organelles between daughter cells (Savage et al., 2010). Syn7942 cytoplasm densely accommodates carboxysomes and chromosomes, which are interspersed with each other (**Figure 3-3A**). It is conceivable that the organisation and dynamics of carboxysomes correlate with the partitioning of chromosomes. On the other hand, the disruption of parA resulted in unequal positioning of carboxysomes, but did not interfere with chromosome organisation, suggesting that the spatial partitioning of carboxysomes and chromosomes in Syn7942 is likely regulated separately (Jain et al., 2012). The detailed underlying mechanism awaits further examination.

Given the spatial constraints in the cytoplasm and the large volume of carboxysomes, the broader distribution of carboxysomes within the cell may suggest the specific associations between carboxysomes and the thylakoid membrane. Indeed, such an interaction has been deduced due to the facts that Rubisco can be found not only in the cytosol, but also near the thylakoid membranes (Agarwal et al., 2009). The structural heterogeneity and dynamics of cyanobacterial thylakoid membranes are fundamental to the physiological regulation of photosynthetic electron transport for energy conversion (Liu, 2016). It is feasible that components in the cycle have specific subcellular positioning to take advantage of the supplied energy and functionally coordinate with each other. Moreover, systems analysis suggested there might be a gradient of CO_2 concentration from the cell membrane to the centre of the cell cytoplasm (Mangan and Brenner, 2014). Thus, changes in the subcellular localisation of carboxysomes, in particular along the short axis of Syn7942 cells, probably render a means for modulating the assimilation of CO_2 within the cell.

Whether there are free Rubisco proteins that are not encapsulated within carboxysomes is an open question. Our confocal imaging did not demonstrate the existence of free Rubisco in the Syn7942 cytoplasm. It was further confirmed by our finding that no visible band of free Rubisco was determined using native PAGE and immunoblot analysis of the soluble fraction. However, it cannot be excluded that the amount of free Rubisco is too low to detect, given the inherent resolution limitations of confocal microscopy and the sensitivity of immunoblot analysis.

The redox state of photosynthetic electron transport chain functions as the key controller of the distribution of respiratory complexes (Liu et al., 2012), photosystem composition (Fujita et al., 1987), photosynthetic state transitions (Mullineaux and Allen, 1990) and the modulation of the circadian clock (Ivleva et al., 2006; Wood et al., 2010). In this study, we report that the redox state of the photosynthetic electron transport chain located in thylakoid membranes has an effect on the subcellular positioning of β-carboxysomes in Syn7942. The widespread and clumping distribution of carboxysomes is determined by the oxidised state of photosynthetic electron transport chain, whereas the linear positioning of carboxysomes along the cell length is ascribed to the reduced state of

photosynthetic electron transport chain (**Figure 3-14**). Nontreated cells present similar clustering organisation of carboxysomes as DCMU-treated cells, indicating that the PQ pool is oxidised upon the illumination condition, which possibly triggers the state transition to "State 1" (Mullineaux and Allen, 1990) or probably due to the high ratio of photosystem I and photosystem II in cyanobacteria (Howitt et al., 2001). Our results further reveal that the reorganisation of carboxysomes appears as a long-term regulation process. It is reminiscent of the previous finding showing that the constrained diffusive dynamics of β-carboxysomes in Syn7942 (Savage et al., 2010).

It has been reported that *de novo* gene expression of DNA replication components in Syn7942 is dependent on the photosynthetic electron transport activity (Ohbayashi et al., 2013). Here, we show that both DCMU and DBMIB can inhibit the accumulation of carboxysome proteins and, thereby, the formation of carboxysomes (**Figure 3-13**). Whether there are indirect effects of protein synthesis on the spatial positioning of carboxysomes needs further characterisation. Our results demonstrate explicitly that DCMU and DBMIB treatments could result in distinct carboxysome positioning in the cell, whereas the carboxysome numbers, and hence probably the expression of carboxysome proteins, are comparable (**Figure 3-14**), suggesting that the effects of electron transport inhibitors on protein synthesis seem not to correlate with the changes in carboxysome positioning.

In addition to the β-carboxysome positioning, redox regulation is also important for the β-carboxysome biosynthesis and function. Carboxysomes may preferably retain an independent redox environment from that of the cytosol, by the semipermeable shell that can selectively exclude the entry of thioredoxin and other redox equivalents into the interior (Pena et al., 2010; Rae et al., 2013). The shell encapsulation allows the establishment of an oxidising micro-environment within the β-carboxysome (Chen et al., 2013). The thioredoxins in the cytoplasm could reduce the redox damage to carboxysome components and enhance the carboxysome biogenesis and maturation (Rae et al., 2013). In addition, the independent redox modulation of the carboxysome lumen was deduced to be vital for the activities of carboxysome enzymes. The oxidising environment could favour the activation of CcaA (Price et al., 1992) and carbonic anhydrase function of CcmM (Pena et al., 2010).

Chapter 3

Cyanobacterial CO_2-concentrating mechanisms (CCMs) comprise carboxysomes, CO_2 uptake complex NDH-1, and HCO_3^- transporters (Price et al., 2008). To date, we have demonstrated that the *in vivo* distributions of carboxysomes (this work) and NDH-1 complexes (Liu et al., 2012) were both regulated by the redox state of photosynthetic electron transport chain, suggesting the potential interplay between the two components. In addition, given that the transcriptional levels of HCO_3^- transporters were also regulated by light through photosynthetic electron flux (Burnap et al., 2013; McGinn. et al., 2004), it is likely that the organisation of entire CCM pathway in cyanobacteria is modulated, in an integrated network context, by the light-mediated redox regulation of photosynthetic electron flow. Therefore, thylakoid membrane re-modelling during environmental adaption might play a role in the regulation of CCM pathway in the cell, which needs to be determined experimentally in future.

Chapter 4

Single-organelle quantification reveals stoichiometric and structural variability of bacterial CO_2-fixing organelles dependent on the environment

Chapter 4

4.1 Introduction

Stoichiometry, the relative quantities of each component in supramolecular complex is crucial for maintenance and regulation of complex functions. In cyanobacteria, the stoichiometry of photosystem and electron transfer proteins were demonstrated to be the key factor that sustained optimised photosynthesis (Ungerer et al., 2018). Regarding carboxysomes, several studies have indicated the strong relevance of assembly structures and functions to the stoichiometry of particular proteins over carboxysomes. Altering the stoichiometry of CcmM isoforms would yield carboxysomes at greater/smaller sizes and strong correlation was found over the relative content of CcaA and CcmM58 in formed carboxysomes (Long et al., 2011). Manipulation of the native stoichiometry through synthetic biology by merging CcmM, CcmN and CcaA into a chimeric protein generated small carboxysomes with one-fold reduced Rubisco content, suggesting the necessity of proper CcmM-CcmN-CcaA stoichiometry in controlling the carboxysome size and enzyme loadings (Gonzalez-Esquer et al., 2015). Synthetic β-carboxysomes in *E. coli* were found to have obscure shapes unlike native counterparts, and incorrect stoichiometry of contents was believed to be one possible reason (Fang et al., 2018). Overall, understanding the physiological composition and assembly principles of carboxysome building blocks is of fundamental importance to gaining a complete picture of carboxysome formation and heterologously engineering and modulating functional CO_2-fixing organelles to supercharge photosynthetic carbon fixation.

Regarding the overall stoichiometry of carboxysomes, previous estimations on the relative carboxysome protein stoichiometry from either the whole cell lysates or the isolated forms, using immunoblot and mass spectrometry, were not completely accurate and relied strongly on sample preparation and treatment as well as the effectiveness and availability of antibodies (Faulkner et al., 2017; Long et al., 2005; Long et al., 2011; Rae et al., 2012). For instance, stoichiometry is only partially available for CcmK2, rbcL, rbcS, CcmM35/58 and CcaA by immunoblotting (Long et al., 2007). Besides, CcmK3 and CcmK4 were never separated from CcmK2 in quantitative comparison due to the sequence homology and their relatively low abundance. For CcmL, no experimental quantification has been done due to its low abundance both in whole cell lysates or isolated

carboxysomes. The putative 12 pentameric CcmL per carboxysome in current model was based on predictions via the structural homology with virus capsid that contains both pentamers and hexamers. CcmN surprisingly was never detected in isolated carboxysome fraction (Faulkner et al., 2017; Long et al., 2005; Long et al., 2011) by mass spectrometry. CcmO predicted to be occupying the edge of facets which account for 10-30% of the shell surface, has never been detected in isolated carboxysome fraction too (Faulkner et al., 2017; Rae et al., 2012). For rbcX, quantification could not be performed due to the low content in purified carboxysome (Rae et al., 2013). Moreover, due to the multiple numbers of carboxysome in cells, the stoichiometry of individual carboxysome cannot be determined effectively through immunoblot, and mass spectrometry approaches.

The adaption and regulatory mechanisms for stoichiometry in eukaryotic organelles are commonly acknowledged. For instance, organelles such as mitochondria and chloroplast can undertake structural reorganisation to optimise their functions under a variety of environmental stimuli (Bartolák-Suki et al., 2017; Dekker and Boekema, 2005; Pernas and Scorrano, 2016). Meanwhile, we lack evidence and comprehensive documentation of possible stoichiometric regulation in carboxysomes. Multiple transcriptomic data of β-carboxysomes in *Synechocystis* PCC6803 have indicated the proportional changes of different carboxysomal genes on expressional levels answering to changes in environmental conditions such as light intensity and CO_2 availability (Eisenhut et al., 2007; Hihara et al., 2001; Wang et al., 2004), shown in **Figure 4-1**. Under increased levels of CO_2 (**Figure 4-1A and B**), general reduction of expression was observed for carboxysomal genes, particularly for long-term adaptation after 24 hours (**Figure 4-1B**). Higher degrees of decrease were found in the shell and structural proteins compared with interior enzyme proteins. While under increased light illumination (**Figure 4-1C)**, general increase of transcription for major shell protein CcmK1, CcmK2, structural protein CcmM and CcmN were observed. Meanwhile lesser increases were observed for enzyme protein RbcL/S and minor shell protein CcmK3/K4. Minimum increases were found among CcmO, CcmP, CcaA and CcmL. The alternations of relative expression might suggest distinctive stoichiometry over formed carboxysomes under different environmental conditions.

Overall, despite the previous findings, the exact stoichiometry of all building components in the functional carboxysome and how carboxysomes manipulate their compositions, organisations and functions to cope with environmental changes remained uncharacterised before this work.

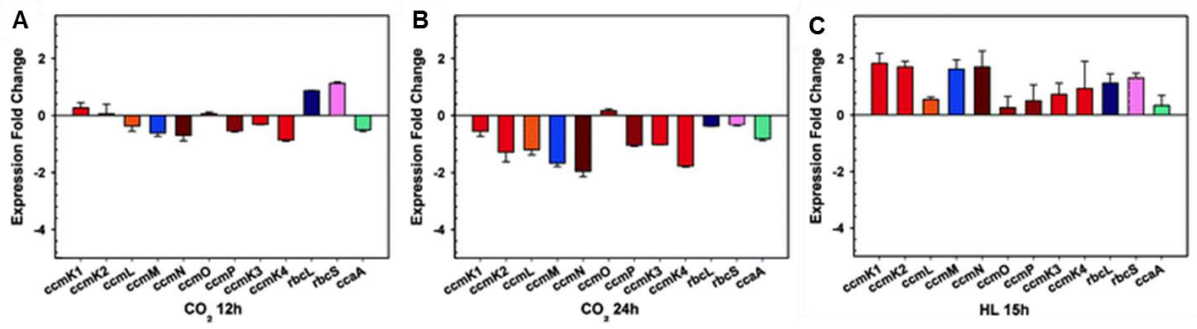


Figure 4-1. Transcriptional regulation of carboxysomal genes from published microarray data in *Synechocystis* PCC 6803 under CO_2 and light regulations. A. 12 hours after switching cells from ambient air to 3 % CO_2 (Wang et al., 2004). **B**. 24 h after switching cells from ambient air to 5 % CO_2 (Eisenhut et al., 2007). **C**. 15 hours after switching cells from 20 to 300 $\mu E \cdot m^{-2}\ s^{-1}$ (HL indicate high light) (Hihara et al., 2001). Figure was modified from (Cai et al., 2012).

Here, we construct a series of Syn7942 mutants with individual integral components of carboxysomes functionally tagged with enhanced yellow fluorescent protein (YFP) and report the *in vivo* characterisation of protein stoichiometry of carboxysomes at the single-organelle level using real-time single-molecule fluorescence microscopy, confocal and electron microscopy, combined with a suite of biochemical and genetic assays. Quantification of the protein stoichiometry of β-carboxysomes in Syn7942 grown under different conditions demonstrates the organisational flexibility of β-carboxysomes, and their ability to modulate functions towards alternations of CO_2 levels and light intensity during cell growth, as well as the regulation of the spatial localisation and mobility of β-carboxysomes in the cell. This study provides fundamental insight into the formation and structural plasticity of carboxysomes and their dynamic organisation towards environmental changes, which could be extended to other BMCs and macromolecular systems. A deeper understanding of the protein composition and structure of carboxysomes will inform strategies for rational design and engineering of functional and adjustable metabolic modules towards biotechnological applications.

4.2 Results

4.2.1 Protein stoichiometry of functional carboxysomes at the single-organelle level

We constructed ten *Synechococcus elongatus* PCC7942 (Syn7942) strains expressing individual β-carboxysome proteins (CcmK3, CcmK4, CcmK2, CcmL, CcmM, CcmN, RbcL, RbcS, CcaA, RbcX) fused with YFP at their C-termini, invividually (**Figure 4-2**). Fluorescence tagging at the native chromosomal locus under the control of their native promoters ensures expression of the fluorescently-tagged proteins in context and at physiological levels (Sun et al., 2016). These YFP-fusion strains, except for RbcL-YFP and CcmK2-YFP, are fully segregated (**Figure 4-2, Figure 4-3**) and exhibit wild-type levels of cell size, growth and carbon fixation within experimental error (**Table 4-1**), consistent with previous observations (Cameron et al., 2013; Faulkner et al., 2017; Savage et al., 2010).

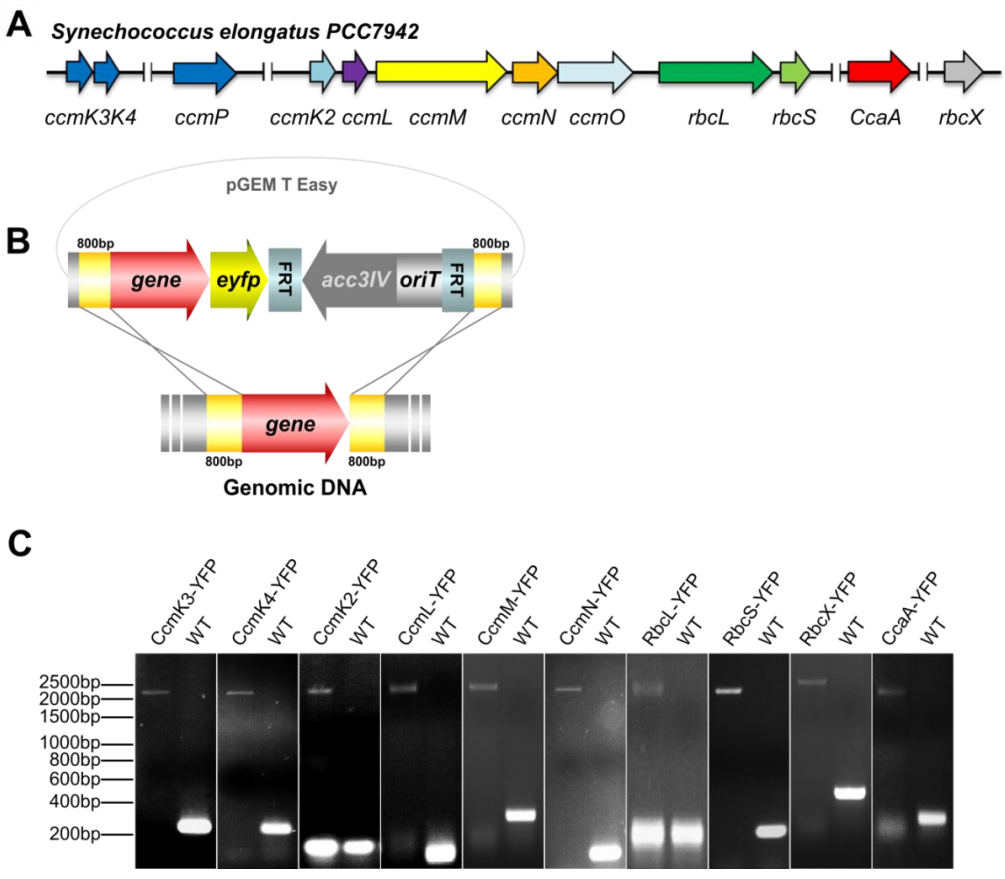


Figure 4-2. Construction and verification of Syn7942 strains with YFP fusion to individual carboxysome proteins. A. Locations of the genes encoding carboxysome proteins in WT Syn7942. **B.** The strategy of YFP fusion using REDIRECT protocol. **C.** PCR verification of the full segregation of YFP-fusion strains, except for CcmK2-YFP and RbcL-YFP which are partially segregated.

Table 4-1. Cell growth, carbon fixation and cell dimension of Syn7942 WT and YFP-fusion mutants under Air/ML. Results were shown as mean ± SD. The sample sizes, as cell numbers, were shown as *n*. No significant differences in growth rate were observed, indicating that YFP tagging has no notable effects on cell growth and physiology.

Strains	**Doubling time (hours) (*n* = 4)**	**Carbon fixation (nmol ·ml^{-1} ·min^{-1}) (*n* = 3)**	**Cell length (μm)**	**Cell width (μm)**	***n***
WT Syn7942	16.82 ± 1.31	3.91 ± 0.08	3.38 ± 1.65	0.88 ± 0.14	1360
CcmK2-YFP	17.26 ± 0.83	3.45 ± 0.09	3.4 ± 1.28	0.82 ± 0.1	951
CcmK3-YFP	18.03 ± 1.22	4.22 ± 1.06	3.66 ± 1.43	0.78 ± 0.08	1015
CcmK4-YFP	17.68 ± 1.44	3.78 ± 0.86	3.6 ± 1.45	0.82 ± 0.08	904
CcmL-YFP	18.59 ± 1.41	3.76 ± 0.47	3.16 ± 1.66	0.82 ± 0.17	668
CcaA-YFP	16.79 ± 0.98	3.38 ± 0.36	3.45 ± 1.39	0.81 ± 0.09	622
CcmN-YFP	17.27 ± 1.08	3.68 ± 0.68	3.34 ± 1.07	0.87 ± 0.07	751
CcmM-YFP	17.74 ± 1.02	3.76 ± 0.47	3.3 ± 1.53	0.86 ± 0.14	705
RbcL-YFP	17.78 ± 1.88	4.01 ± 0.62	3.22 ± 1.19	0.85 ± 0.1	556
RbcS-YFP	18.33 ± 2.13	3.70 ± 1.01	3.03 ± 1.07	0.78 ± 0.09	236
RbcX-YFP	18.14 ± 0.77	3.67 ± 0.87	3.67 ± 1.52	0.82 ± 0.12	799

Also, we also generated RbcL-YFP and CcmK2-YFP strains, which were only partially segregated. Through immunoblot analysis using anti-fluorescence protein, anti-RbcL and anti-CcmK2 antibodies (**Figure 4-3**), we estimate that 29.2 ± 7.1 % (*n* = 4) of total RbcL and 6.0 ± 0.7 % (*n* = 3) of total CcmK2 were tagged with YFP in RbcL-YFP and CcmK2-YFP strains. Nevertheless, we excluded the stoichiometric quantification of RbcL and CcmK2 in this study, given the partial segregation which could result in quantification inaccuracy.

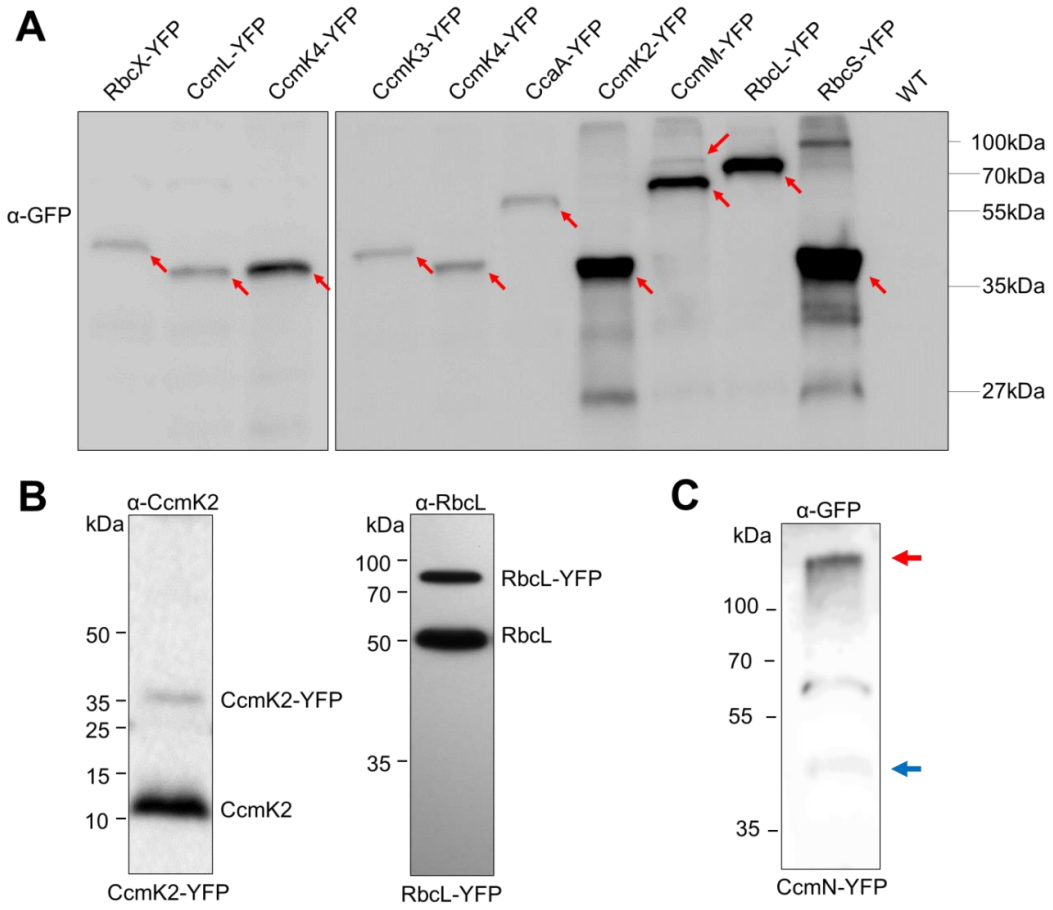


Figure 4-3. Immunoblot analysis of the YFP-tagged Syn7942 strains using the anti-GFP, anti-RbcL and anti-CcmK antibodies of soluble fractions in this study based on SDS-PAGE. A. Immunoblot of soluble fractions of YFP-tagged strains using an anti-fluorescent protein antibody shows tagging of YFP to each carboxysome protein (red arrows). Note that CcmM-YFP shows two bands at 85 and 62kDa, corresponding to CcmM58 and CcmM35, respectively. Samples with the same total protein concentration were loaded to relatively quantify the carboxysome protein content in the soluble fraction (**Table 4-2**). **B.** Immunoblot of soluble fractions of the partially-segregated CcmK2-YFP and RbcL-YFP strains, using anti-CcmK2 and anti-RbcL antibodies. It indicates that 29.2 ± 7.1 % (n = 4) of total RbcL and 6.0 ± 0.7 % (n = 3) of total CcmK2 proteins were tagged with YFP. **C.** Immunoblot of soluble fractions of the CcmN-YFP strain using an anti-fluorescent protein antibody, showing the presence of large aggregates of CcmN that failed to run into the gel, marked by the red arrow. Minor fractions of CcmN that run into the gel are marked by the blue arrow. For all western results, at least 3 independent experiments were done for statistical analysis.

Table 4-2. Carboxysomal protein stoichiometry in cell lysate detected by immunoblotting. Quantification of CcmO and CcmN was not included, as they could form large aggregations in SDS-PAGE and their contents could not be accurately estimated. Data are represented as mean ± SD from at least 3 biological repeats.

Category	Protein	Relative content (%)
Shell proteins	CcmK2	70.4 ± 7.1
	CcmK3	1.3 ± 0.1
	CcmK4	2.0 ± 0.2
	CcmL	0.2 ± 0.01
Structural proteins	CcmM	2.5 ± 0.5
CA enzymes	CcaA	1.7 ± 0.4
Rubisco enzymes	RbcL	12.6 ± 3.5
	RbcS	9.1 ± 0.8
Rubisco chaperone	RbcX	0.1 ± 0.01

We used single-molecule Slimfield microscopy (Plank et al., 2009) to visualise individual carboxysomes that were fused with YFP (**Figure 4-4**, **Figure 4-5**). This technique allows detection of fluorescently-labelled proteins with millisecond sampling, enabling real-time tracking of rapid protein dynamics inside living cells, exploited previously to study functional proteins involved in bacterial DNA replication and remodelling (Badrinarayanan et al., 2012; Reyes-Lamothe et al., 2010), gene regulation in budding yeast cells (Leake, 2018; Wollman et al., 2017), bacterial cell division (Lund et al., 2018), and chemokine signalling in lymph nodes (Miller et al., 2018). Our prior measurements using relatively fast maturing fluorescent proteins such as YFP suggest that less than 15% of fluorescent protein is likely to be in a non-fluorescent immature state (Leake et al., 2008; Shashkova et al., 2018).

Figure 4-4A shows the Slimfield images of three representative Syn7942 strains RbcS-YFP, CcmK4-YFP and CcmM-YFP that grow under ambient air and moderate light (hereafter denoted Air/ML), to determine the protein stoichiometry from different carboxysome structural domains. Single carboxysomes are detected as distinct fluorescent foci in cells of the YFP-fused strains (**Figure 4-4A**, **Figure 4-5**), whose sigma width is approximately 250 nm (n = 100), comparable to the diffraction-

limited point spread function width of our imaging system. We use the number of YFP molecules per fluorescent foci as an indicator of the stoichiometry of the fluorescently-labelled protein subunits in each carboxysome, which we determined by quantifying step-wise photobleaching of the fluorescent tag (Leake et al., 2006) during the Slimfield laser excitation process (**Figure 4-4B to C**) using a combination of Fourier spectral analysis and edge-preserving filtration of the raw data (Leake et al., 2003; Leake et al., 2004). The resulting broad distributions, rendered as a kernel density estimates of protein stoichiometry suggest a variable content of individual components per carboxysome (**Figure 4-4D**), indicative of the structural heterogeneity of β-carboxysomes. The modal average stoichiometry of each protein subunit per carboxysome was defined by the measured peak from each distribution of the raw stoichiometric data (**Figure 4-4D**, **Figure 4-5**), after subtracting the background fluorescence distribution, primarily from chlorophylls, which was determined from the WT cells (**Figure 4-6**).

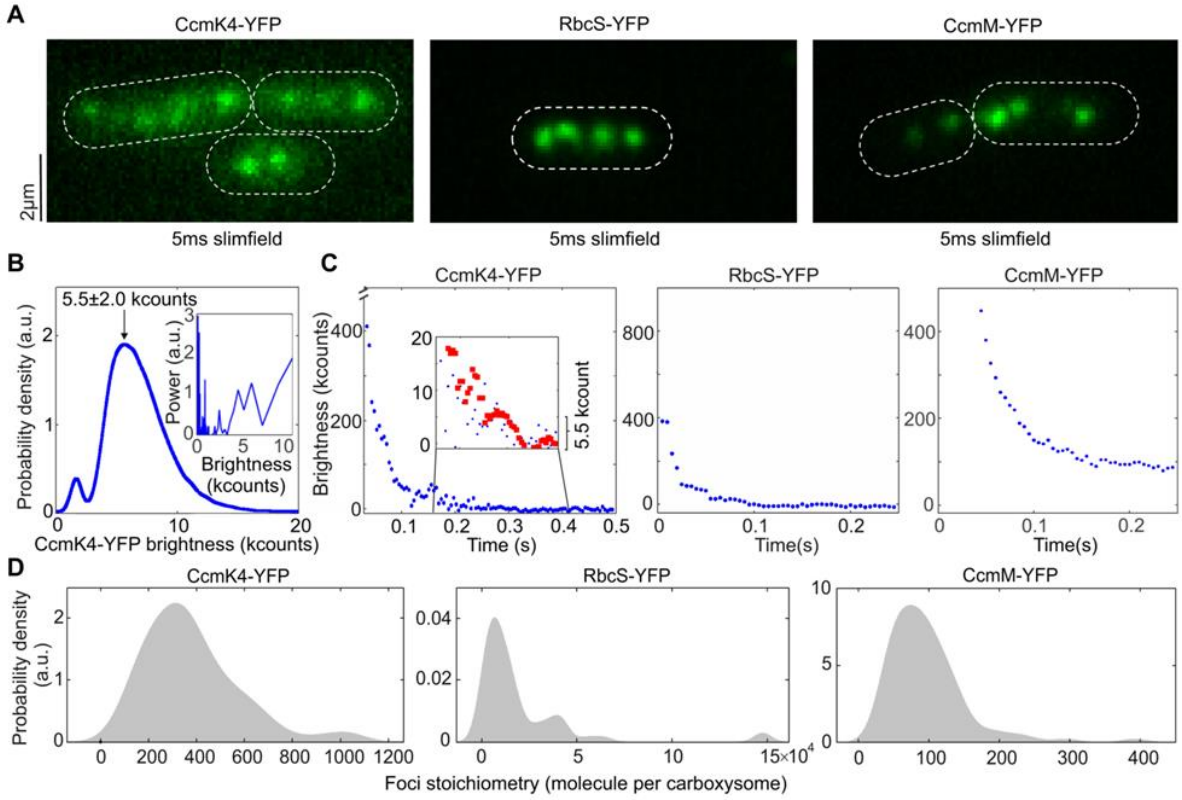


Figure 4-4. Slimfield quantification of cells grown under ambient air/moderate light Air/ML condition. A. Averaged Slimfield images of YFP fluorescence (green) over five frames of strains expressing shell component CcmK4-YFP, the interior enzyme RbcS-YFP and the shell-interior linker protein CcmM-YFP, cell body outlines indicated (white dash line). Scale bar indicates 2 μm. **B.** Distribution of automatically detected foci intensity from the end of the photobleaching,

corresponding to the characteristic intensity of *in vivo* YFP. Inset shows the Fourier spectrum of 'overtracked' foci, tracked beyond photobleaching, showing a peak at the characteristic intensity. **C.** Representative fluorescence photobleaching tracked at ultra-fast speed. The CcmK4 plot shows an inset 'zoomed in' on lower intensity range with step-preserving Chung-Kennedy filtered data in red, with individual photobleaching steps visible at the characteristic intensity. Brightness (kcounts), counts measured per camera pixel multiplied by 1,000. **D.** Distribution of YFP copy number detected for individual carboxysomes in corresponding mutant strains, rendered as kernel density estimates using standard kernel width. Heterogeneity of contents was observed. Also, a "preferable" copy number, represented by kernel density peak values could be determined. Statistics of copy numbers (Peak value ± HWHM) are listed in **Table 4-3** for ML conditions. The corresponding Slimfield images and histogram for complete strain sets were shown in **Figure 4-5**.

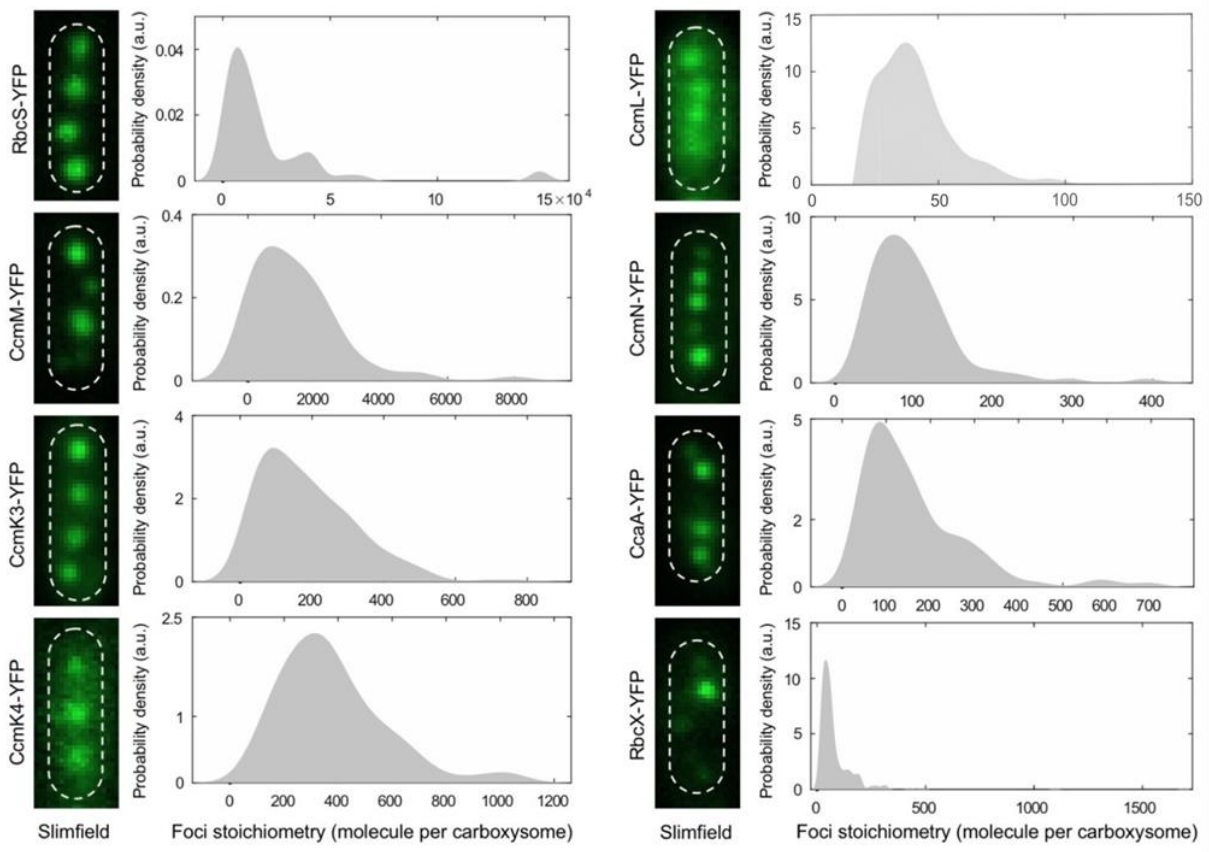


Figure 4-5. Slimfield images of YFP-fusion cells under Air/ML and stoichiometric histogram of copies of YFP per carboxysome. Fluorescence foci (green) indicate individual carboxysomes. Cell borders are outlined by white dash lines. In the histograms, the major peaks were acquired by kernel density fitting, representing the copy numbers of each protein per carboxysome. Sample sizes for individual strains are 60 (RbcS), 219 (CcmK3), 77 (CcmK4), 316 (CcmL), 71 (CcmM), 86 (CcmN), 95 (CcaA) and 211 (RbcX) respectively.

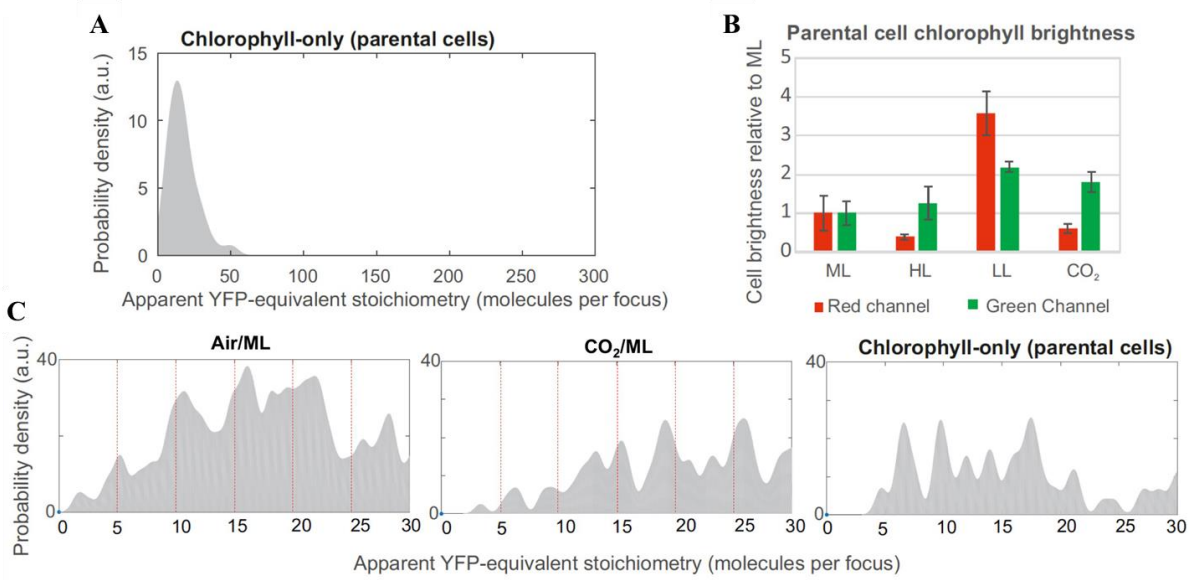


Figure 4-6. Normalisation of chlorophyll during Slimfield imaging for Syn7942 strains. A. The apparent YFP-equivalent stoichiometry of Chlorophyll estimated from Wild-type Syn7942 at Air/ML condition. **B.** Relative brightness of chlorophyll detected by Slimfield imaging in the YFP (green) channel and chlorophyll (red) channel, showing expected trends between growth conditions in the red channel but no clear trend in the green channel, such that the Air/ML distribution in **A** was used to correct all growth condition. **C.** Uncorrected kernel density estimates for CcmL stoichiometry in Air/ML and CO_2/ML growth conditions and apparent stoichiometry of chlorophyll, generated with a kernel width of 0.5 YFPs. This low kernel width allows periodic features to be seen, in this case showing 5mer periodicity in CcmL stoichiometries.

We estimated the stoichiometry of building proteins in single carboxysomes grown under Air/ML through a cell-by-cell based Slimfield imaging using numerical integration of pixel intensities (Wollman and Leake, 2015) (**Table 4-3**) in each carboxysome divided by the intensity of a single eYFP (Figure 4-4B) – full details in Chapter 2. Rubisco appears to be the predominant components of the β-carboxysome, as indicated by the RbcS content. CcmM is the second abundant elements; there are over 700 copies of CcmM per β-carboxysome. In addition, CcmK4 content is greater than that of CcmK3 by a factor of 3.8. CcmL, CcmN, CcaA and RbcX are the minor components in the β-carboxysome. Our results reveal that there are 37 CcmL subunits per carboxysome, with the raw stoichiometry distribution showing some indications of peaks at multiples of ~ five molecules indicative of multiples of CcmL pentamers (**Figure 4-6C**), consistent with the atomic structure of CcmL (Tanaka et al., 2008). A modal average of 37 CcmL molecules, therefore, suggests that a single

carboxysome contains an average of 7.4 CcmL pentamers, less than 12 CcmL pentamers that are postulated to occupy all the vertices of the icosahedral shell (Bobik et al., 2015; Kerfeld et al., 2018). It is feasible that not all vertices of the carboxysome structure are capped by CcmL pentamers, as BMC shells deficient in pentamers could still be formed without notable structural variations (Cai et al., 2009; Hagen et al., 2018; Lassila et al., 2014). To our knowledge, this is the first characterisation of protein stoichiometry at the level of single functional carboxysomes in their native cellular environment.

Table 4-3. Protein stoichiometry of the carboxysome and its variability in Syn7942 grown under Air/ML, CO_2/ML, LL and HL conditions determined from Slimfield and confocal microscopy imaging. Stoichiometry is presented as Peak value ± HWHM, and the sample sizes are indicated as *n*. Peak values were determined from the Slimfield stoichiometry profiles of each carboxysome proteins (**Figure 4-4**, **Figure 4-5**). Quantification of the CcmL under the four conditions was acquired from Slimfield imaging for accurate measurement of the copies of shell pentamers for capping the carboxysome structure. Copies of other carboxysome proteins were calculated using Slimfield results (grey) with definitive counts of protein copies under Air/ML (**Figure 4-5**) in combination with relative quantification of each protein under the four conditions from confocal imaging (**Figure 4-9** and **Figure 4-10**). The structures of protein were documented in (Kerfeld et al., 2005; Long et al., 2007; Tanaka et al., 2007; Tanaka et al., 2008; Long et al., 2011; Kinney et al., 2012; McGurn et al., 2016). *CcmM58 have trimer formation, the Monomer was designated to the majority of CcmM35.

Category	Structure	Protein	Air/ML		CO_2/ML		LL		HL	
			Peak value ± HWHM	Number of functional units	Peak value ± HWHM	Number of functional units	Peak value ± HWHM	Number of functional units	Peak value ± HWHM	Number of functional units
Shell proteins	Hexamer	CcmK3	92 ± 148 (*n* = 219)	15 ± 25	172 ± 83 (*n* = 2048)	29 ± 14	83 ± 31 (*n* = 1516)	14 ± 5	87 ± 52 (*n* = 2155)	14 ± 9
		CcmK4	314 ± 194 (*n* = 77)	52 ± 32	562 ± 263 (*n* = 1918)	94 ± 44	313 ± 121 (*n* = 1766)	52 ± 20	304 ± 95 (*n* = 3215)	51 ± 16
	Pentamer	CcmL	37 ± 17 (*n* = 316)	7.4 ± 3.4	66 ± 24 (*n* = 311)	13.2 ± 4.8	34 ± 15 (*n* = 394)	6.8 ± 3.0	69 ± 24 (*n* = 220)	13.8 ± 4.8
Structural proteins	Monomer*	CcmM	719 ± 1433 (*n* = 71)	719 ± 1433	468 ± 425 (*n* = 2313)	468 ± 425	483 ± 366 (*n* = 3655)	483 ± 366	1176 ± 691 (*n* = 2318)	1176 ± 691
	Monomer	CcmN	74 ± 51 (*n* = 86)	74 ± 51	52 ± 28 (*n* = 3143)	52 ± 28	51 ± 20 (*n* = 4022)	51 ± 20	82 ± 34 (*n* = 5074)	82 ± 34
CA enzyme	Hexamer	CcaA	86 ± 81 (*n* = 95)	14 ± 14	129 ± 86 (*n* = 1354)	21 ± 14	65 ± 21 (*n* = 217)	11 ± 4	122 ± 59 (*n* = 2837)	20 ± 10
Rubisco enzyme	L_8S_8	RbcS	6822 ± 9200 (*n* = 60)	853 ± 1150	4401 ± 6655 (*n* = 894)	550 ± 832	2934 ± 5492 (*n* = 752)	367 ± 687	12057 ± 5186 (*n* = 1974)	1507 ± 648
Rubisco chaperone	Dimer	RbcX	39 ± 32 (*n* = 211)	20 ± 16	38 ± 10 (*n* = 1370)	19 ± 5	40 ± 9 (*n* = 1402)	20 ± 5	40 ± 9 (*n* = 1861)	20 ± 5

As a control, we fused RbcL with mYPet, a monomeric-optimised variant of YFP. The RbcL-YFP and RbcL-mYPet cells show no significant difference in the subcellular distribution of carboxysomes as well as cell doubling times and carbon fixation (**Figure 4-7**), demonstrating that there are no measurable artefacts due to putative effects of dimerisation of the YFP tag.

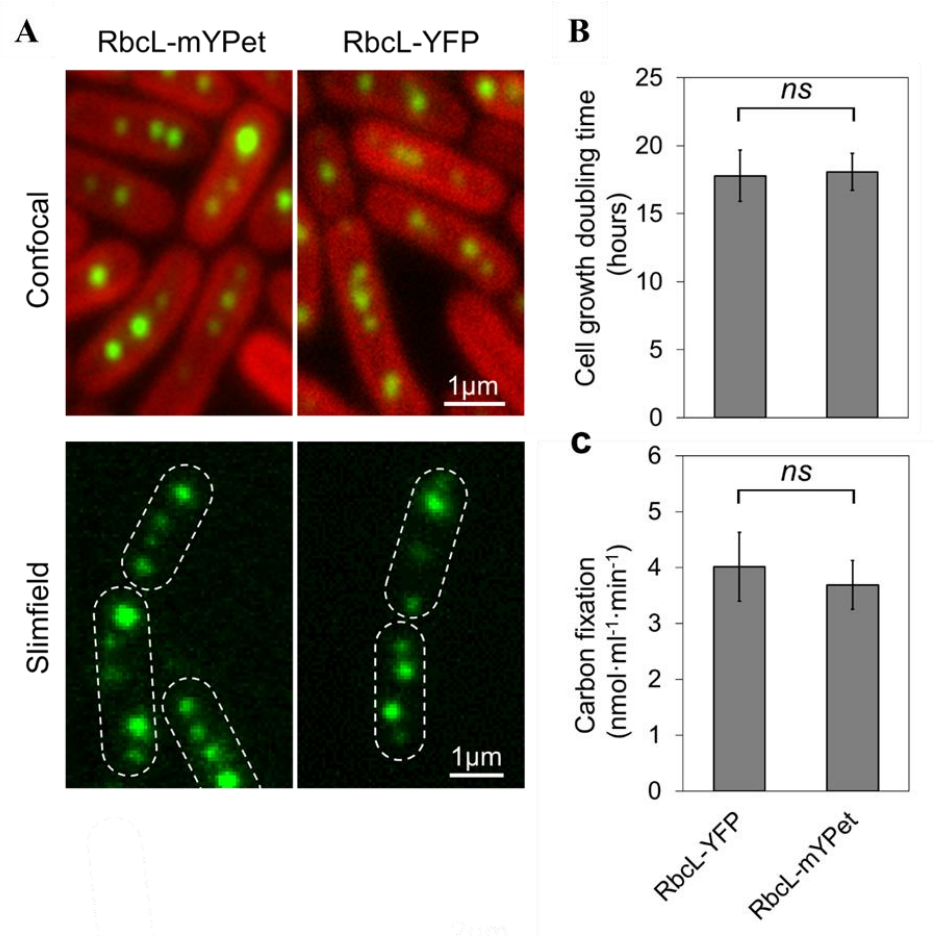


Figure 4-7. Comparison of YFP and mYPet tagging to RbcL in Syn7942. A. Confocal and Slimfield microscopy images for RbcL-mYPet and RbcL-YFP. No significant effects were detected on the doubling time **B** (p = 0.82, n = 4, student t-test) and carbon fixation capacity **C** (p = 0.50, n = 3) of RbcL-mYPet and RbcL-YFP strains. No significant differences in RbcL stoichiometry, cell growth and carbon fixation, suggesting that there are no measurable artefacts due to putative effects of dimerisation of the YFP tag.

We also examined the relative abundance of individual carboxysome proteins in the YFP-fusion Syn7942 strains in cell lysates, using immunoblot probing with an anti-fluorescent protein antibody (**Figure 4-3A**, **Table 4-2**). To compare with the stoichiometry obtained from Slimfield imaging, we normalised the abundance of carboxysome proteins estimated from immunoblot analysis, using the

RbcS content per carboxysome determined by Slimfield imaging. It appears that the content of β-carboxysome proteins determined by immunoblot is generally greater than that within the carboxysome characterised by Slimfield. Despite the potential effects caused by YFP fusion, it could suggest the presence of a "storage pool" of carboxysome proteins located in the cytoplasm, which are involved in the biogenesis, maturation and turnover of carboxysomes. The ratio of RbcL/S is 8:5.8 (*n* = 4) (**Table 4-2**), in line with the previously proposed $RbcL_8S_5$CcmM35 assembly of Rubisco complexes in Syn7942 (Long et al., 2011).

4.2.2 Stoichiometry of carboxysome proteins exhibit a dependence on the microenvironment conditions of live cells

Our previous study showed that the content and spatial positioning of β-carboxysomes in Syn7942 are dependent upon light intensity during cell growth, revealing the physiological regulation of carboxysome biosynthesis (Sun et al., 2016). Whether the stoichiometry of different components in the carboxysome structure changes in response to fluctuations in environmental conditions is unknown. Here we addressed this question by taking advantage of the far greater throughput of confocal microscopy compared to Slimfield, while still using the single-molecule precise Slimfield data as a calibration to convert the intensity of detected foci from confocal images into estimates for absolute numbers of stoichiometry. We achieved this by identifying the peak value of the foci intensity distribution from each given cell strain obtained from confocal imaging with the peak value of the measured Slimfield foci stoichiometry distribution for the equivalent cell strain under Air/ML condition. This approach allows us to generate a conversion factor which we then applied to subsequent confocal data acquired under lower light (LL), higher light (HL) and ML with air supplemented by 3% CO_2, and to estimate relative changes in the stoichiometry of carboxysome building components using relatively large numbers of cells, without the need to obtain separate Slimfield data sets for each condition (**Figure 4-8** to **Figure 4-10**).

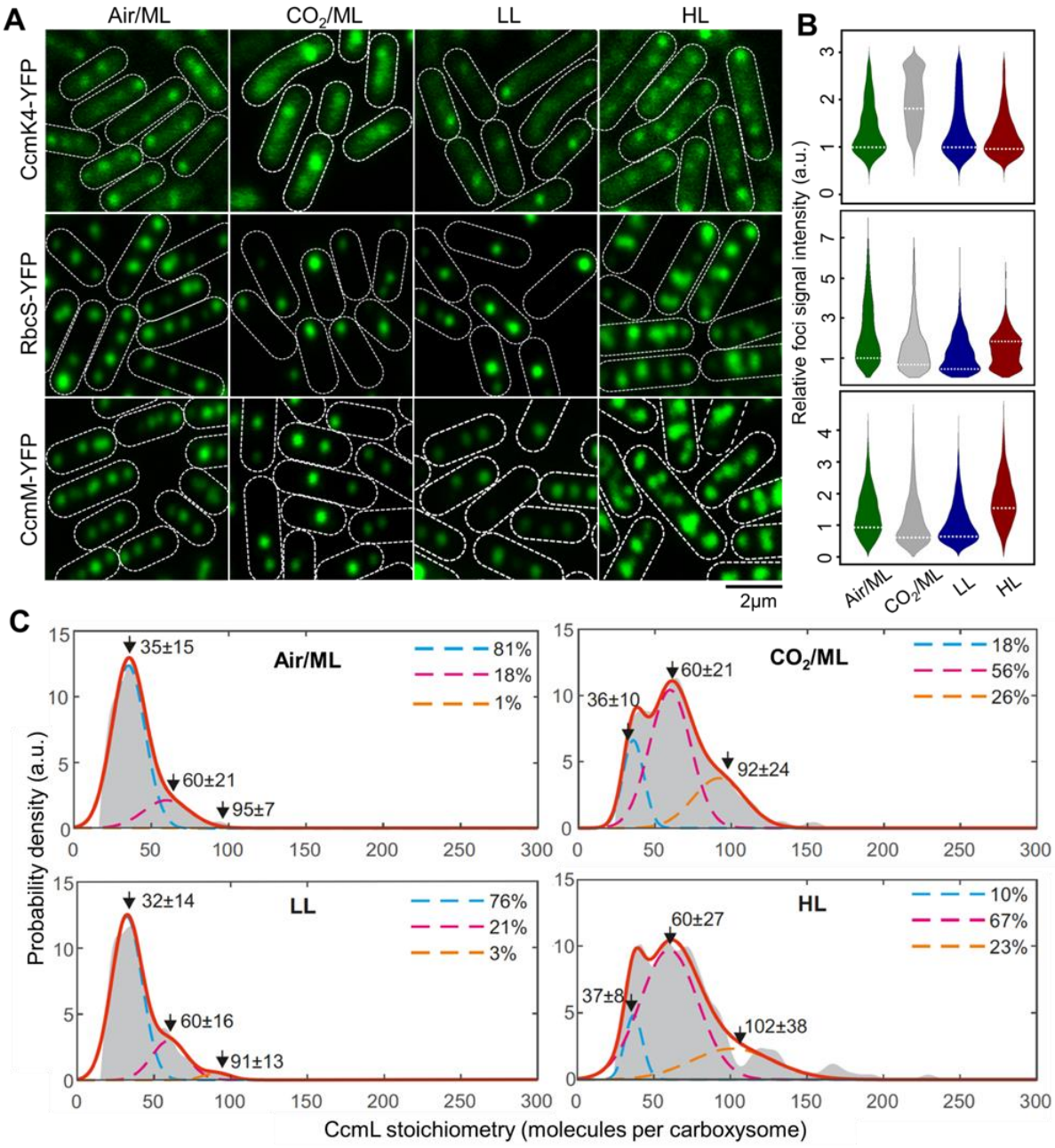


Figure 4-8. Relative protein quantification of CcmK4, RbcS and CcmM in the carboxysome under different CO_2 levels and light intensities using confocal microscopy. A. Confocal images of CcmK4-YFP, RbcS-YFP and CcmM-YFP strains under Air/ML, CO_2/ML, LL and HL. Fluorescence foci (green) indicate carboxysomes and cell borders were outlined by white dashed lines. **B.** Violin plot of carboxysome intensities under Air/ML, CO_2/ML, LL and HL, normalised to kernel density ML peak values (peaks marked by white dash lines). **C.** Kernel density estimates of CcmL carboxysome copy number grown under Air/ML, CO_2, LL and HL detected by Slimfield and corrected for chlorophyll. Triple Gaussian fits are indicated as coloured dashed lines with summed fit in red. The percentage of each Gaussian is indicated aside.

Figure 4-8A shows confocal fluorescence images of CcmK4-YFP, RbcS-YFP, and CcmM-YFP strains grown under Air/ML, 3% CO_2 (CO_2/ML), LL and HL. The confocal images reveal patterns of cellular localisation of carboxysomes similar to those observed with Slimfield microscopy (**Figure 4-5**).

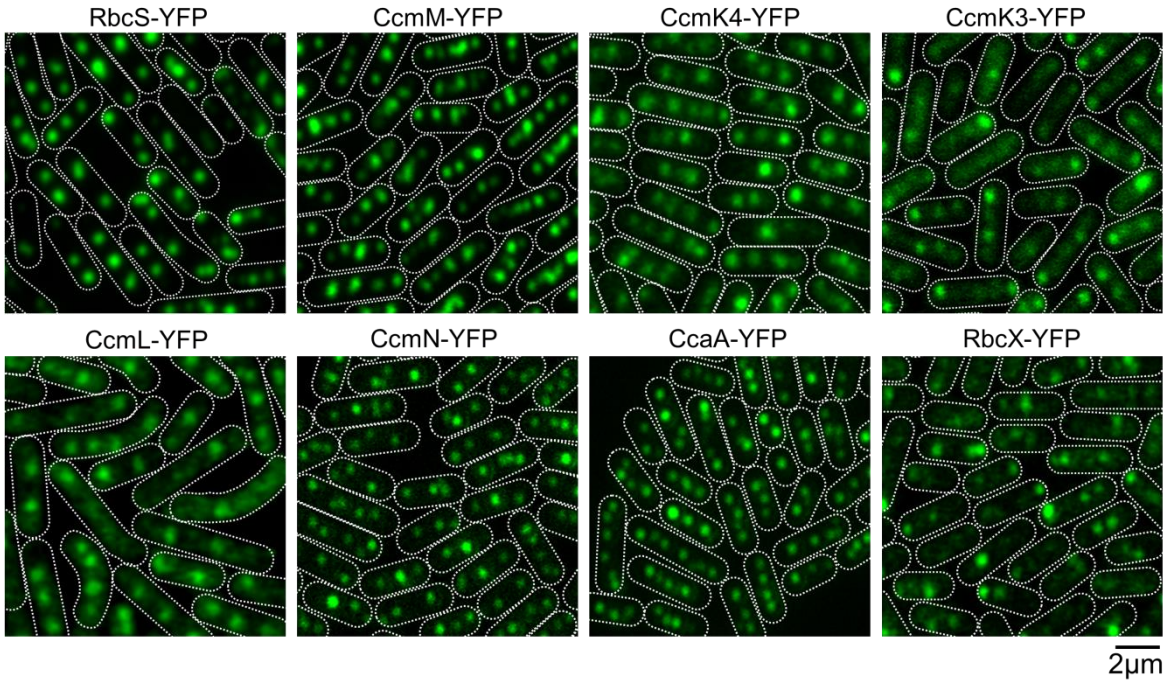


Figure 4-9. Confocal images of YFP-tagged cells. YFP signals are shown in green, indicative of YFP-labeled carboxysomes. White dash lines indicate cell borders. (scare bar = 2 μm).

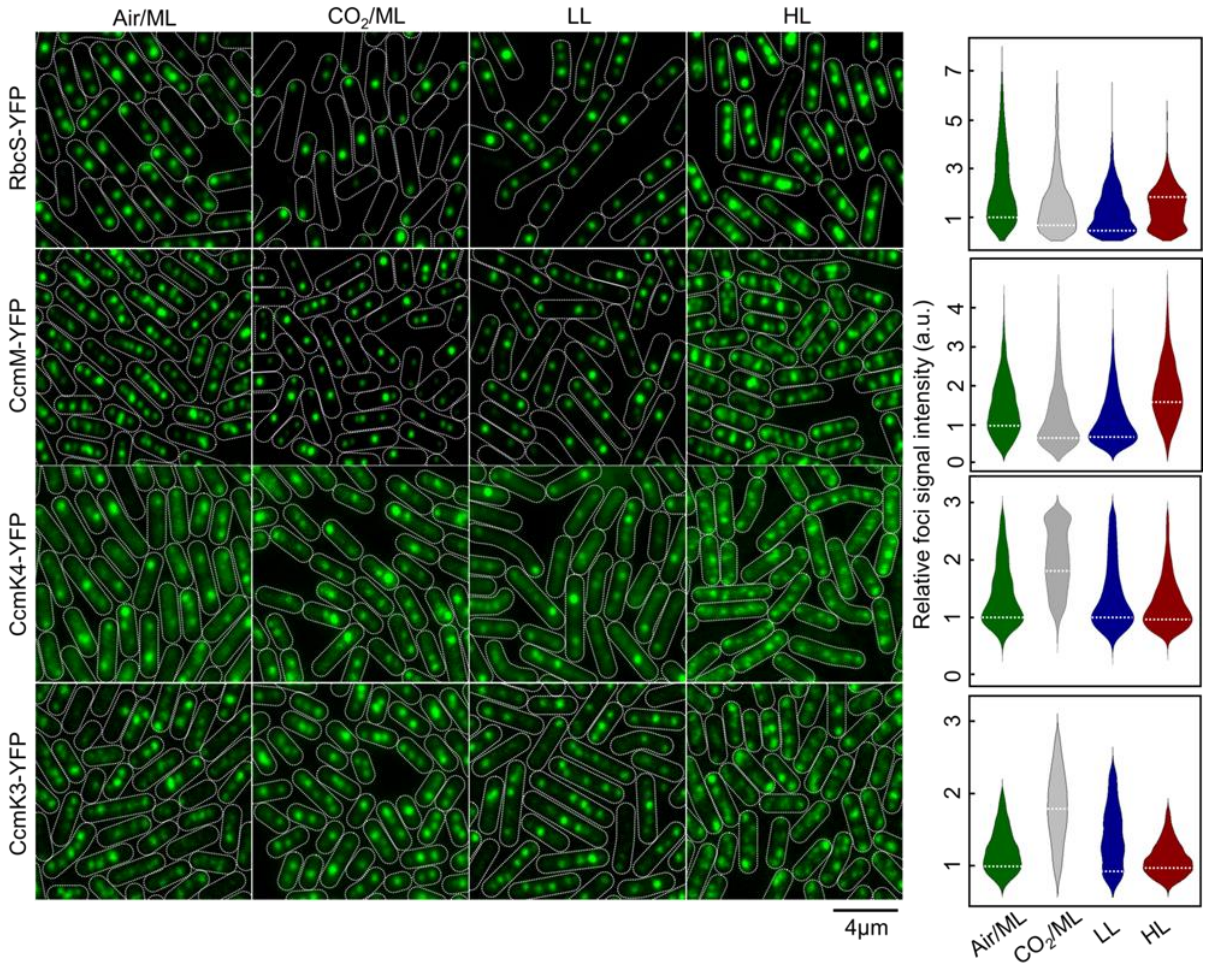


Figure 4-10. Confocal images of RbcS-YFP, CcmM-YFP, CcmK4-YFP and CcmK3-YFP cells under Air/ML, CO_2, LL, and HL and distribution profiles of carboxysome protein signal intensity. Imaging on the same strain under the four treatments was performed with the same parameters, allowing direct comparison of fluorescence intensities. Carboxysome and cell borders (shown as green and white dash lines, respectively) were optimised for presentation. The peak values obtained from kernel density fitting of the intensity distribution, as shown by white dashed lines, were used for examining changes in the abundance of the same carboxysome proteins under different conditions. Scale bar = 4 μm.

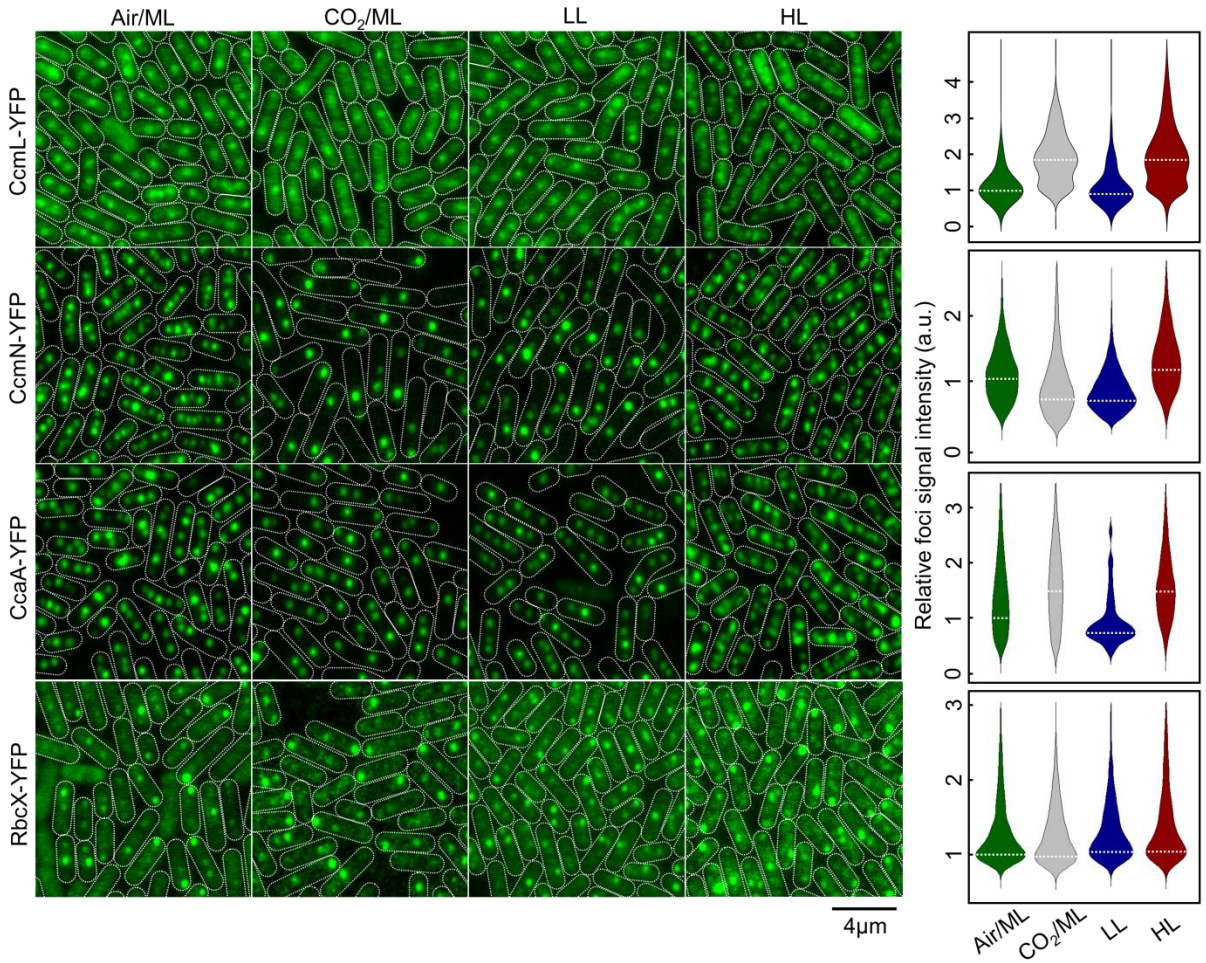


Figure 4-11. Confocal images of CcmL-YFP, CcmN-YFP, CcaA-YFP and RbcX-YFP cells under Air/ML, CO_2, LL, and HL and distribution profiles of carboxysome protein signal intensity (continuing **Figure 4-10**).

We analysed the confocal images to detect carboxysome fluorescent foci within the cells and quantify their fluorescence intensities (**Figure 4-8B**, **Figure 4-10** and **Figure 4-11**). We find that the number of carboxysomes per cell is dependent on the growth condition: it was reduced under CO_2/ML in contrast to Air/ML, whereas HL increases the abundance of β-carboxysomes, consistent with previous findings (Sun et al., 2016; Whitehead et al., 2014). As a common feature, the abundance of all the proteins in the β-carboxysome is modulated under distinct growth conditions. For instance, both RbcS and CcmM have a higher content per carboxysome under HL compared with that under other conditions, whereas the CcmK4 content per β-carboxysome increase under 3% CO_2 (**Figure 4-8B**). The dependence of carboxysome protein stoichiometry inferred from the peak values of the

stoichiometry distributions under different cellular microenvironmental conditions is summarised in **Table 4-2**.

Interestingly, we find that the variation of CcmL abundance per carboxysome rises with increasing light illumination and CO_2 availability (**Figure 4-8C**). The measured stoichiometry distribution of CcmL pentamers suggests the presence of three populations: (I) carboxysomes with < 60 CcmL subunits (in the range of 32-37); (II) carboxysomes with 60 CcmL subunits, consistent with the expectation that 12 vertices of the icosahedral carboxysome are fully occupied by CcmL pentamers (Kerfeld et al., 2018; Rae et al., 2013; Tanaka et al., 2008); (III) carboxysomes with > 60 CcmL subunits (in the range of 91-102). Using a nearest-neighbour model to estimate the probability for the diffraction-limited optical images of individual carboxysomes in a cell, we find that the Population III carboxysomes represent random overlap of two or more carboxysomes from the Population I and II (**Figure 4-8C**). Population I represent a "non-complete capped" state in which not all vertices in the icosahedron are occupied by CcmL pentamers. We find the characteristic stoichiometry of the Population I carboxysomes increases with the enhancement of light intensity during cell growth, from 32 CcmL molecules (LL) to 35 (ML) and 37 (HL), with HL have a significantly smaller proportion (23 %) of "non-complete capped" carboxysomes compared to ~80% under LL and ML conditions. Supplementing the air with 3% CO_2 under ML similarly resulted in a substantial decrease in the proportion of "non-complete capped" carboxysomes in the population (18 %) comparable to the HL condition in the absence of any supplemental CO_2. These findings suggest a dependence of carboxysome assembly which may allow adaptation towards microenvironmental changes, i.e. the increase in the population of capped carboxysomes in situations which are favourable towards photosynthesis (HL conditions and locally-raised levels of CO_2).

This finding is also validated by the changes in protein abundance of other carboxysome components under environmental regulation (**Table 4-2**, **Figure 4-10** and **Figure 4-11**). Variations of protein content in carboxysomes under CO_2/ML vs Air/ML, and HL vs LL conditions, likely dependent on the locations of corresponding genes in the genome, indicate distinct fashions of stoichiometric regulation

of carboxysome building blocks (**Figure 4-12**). The abundance of CcmK3 and CcmK4, whose encoding genes are distant from the *ccmKLMNO* operon (Sommer et al., 2017), increases under 3 % CO_2 and remain relatively constant under HL/LL, contrary to the changes in the abundance of CcmN, and CcmM located in the *ccm* operon. In addition, the ratio of CcmK4: CcmK3 per carboxysome appear to be relatively constant, in the range of 3.6−4.1 (**Table 4-4**), indicating the organisational correlation between CcmK3 and CcmK4 within the β-carboxysome structure. The content of the putative Rubisco chaperone RbcX in each carboxysome remains unaltered under different conditions, probably ascribed to the fact that its encoding gene is distant from the *Rubisco* and *ccm* operons in Syn7942.

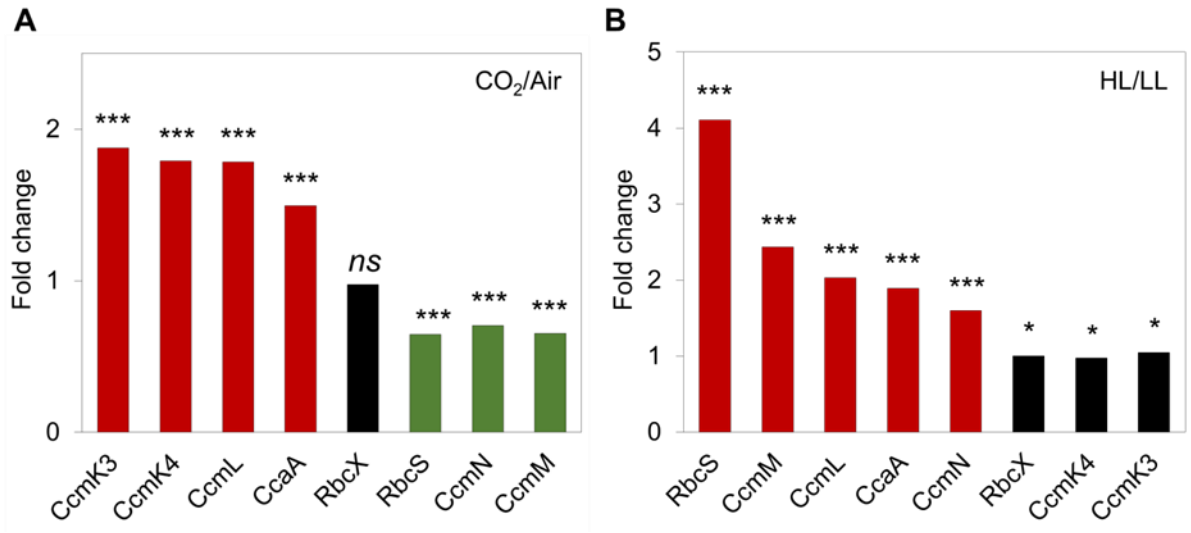


Figure 4-12. Changes in carboxysome protein stoichiometry by increase in CO_2 levels and light intensity. A. Comparison of carboxysome protein stoichiometry under CO_2 treatment. Increase in the CO_2 concentration resulted in the rise of CcmK3, CcmK4, CcaA and CcmL contents and the decline of RbcS, CcmN and CcmM contents. **B.** Comparison of carboxysome protein stoichiometry under light intensity treatment. Increased light intensity led to the elevation of RbcS, CcmM, CcmL, CcaA and CcmN contents, whereas the abundance of RbcX, CcmK3 and CcmK4 contents per carboxysome do not change dramatically. The Peak value ratios of each protein under paired conditions were displayed as fold changes. Man-Whitney U-tests were performed comparing numbers of functional units of target proteins from CO_2/ML to Air/ML and HL to LL separately, P-values < 0.05, <0.005 >0.05 were marked with *, *** and ns respectively.

4.2.3 Variation of carboxysome diameter represents a strategy for manipulating carboxysome activity to adapt to environmental conditions

The change in the protein content per carboxysome signifies the variation of β-carboxysome size and organisation among different cell growth conditions. Indeed, electron microscopy of WT Syn7942 cells substantiated the variable structures of β-carboxysomes in response to changing environment (**Figure 4-13A** and **B**). The average diameter of β-carboxysomes is 192 ± 41 nm (mean ± SD, *n* = 33) in Air/ML, 144 ± 24 nm (*n* = 25) in 3% CO_2, 151 ± 22 nm (*n* = 27) in LL, and 208 ± 28 nm (*n* = 51) in HL (**Figure 4-13B**, **Table 4-4**, **Figure 4-14**). These results reveal that both the CO_2 level and light intensity can result in changes in carboxysome size (**Figure 4-13B**). Larger β-carboxysomes can accommodate more Rubisco enzymes (estimated by RbcS content) (**Figure 4-13C**). An exception is the carboxysomes under LL, which are around 5% larger than the carboxysomes under 3% CO_2 but comprises only 67 % of Rubisco per carboxysome under CO_2 (**Figure 4-13C**, **Table 4-4**). EM images reveal that the lumen of β-carboxysomes synthesised under LL often contain regions with low protein density (**Figure 4-13A**, arrows; **Figure 4-14**), likely resulting in the reduced and uneven Rubisco loading within the β-carboxysome.

We also find that CO_2-fixing activity per carboxysome increases as the β-carboxysome structure enlarges correlated to strong light intensity during cell growth (**Figure 4-13D**), demonstrating the correlation between β-carboxysome structure and function *in vivo*. Moreover, CO_2-fixation activity per Rubisco of the β-carboxysome declines under HL as the carboxysome size (**Figure 4-13E**) and Rubisco density in the carboxysome lumen increase (**Table 4-4**). This may suggest that Rubisco density and local Rubisco interactions are important for determining CO_2-fixation activity of individual Rubisco (**Table 4-4**). Interestingly, the relatively small β-carboxysomes under 3% CO_2 exhibit high CO_2-fixing activities per Rubisco and per carboxysome, compared with β-carboxysomes under other conditions. The enhanced carbon fixation capacity under 3% CO_2 might be correlated with the increase in CcmK3 and CcmK4 content (**Figure 4-12**), as it has been shown that depletion of CcmK3/CcmK4 impedes carbon fixation of carboxysomes (Rae et al., 2012).

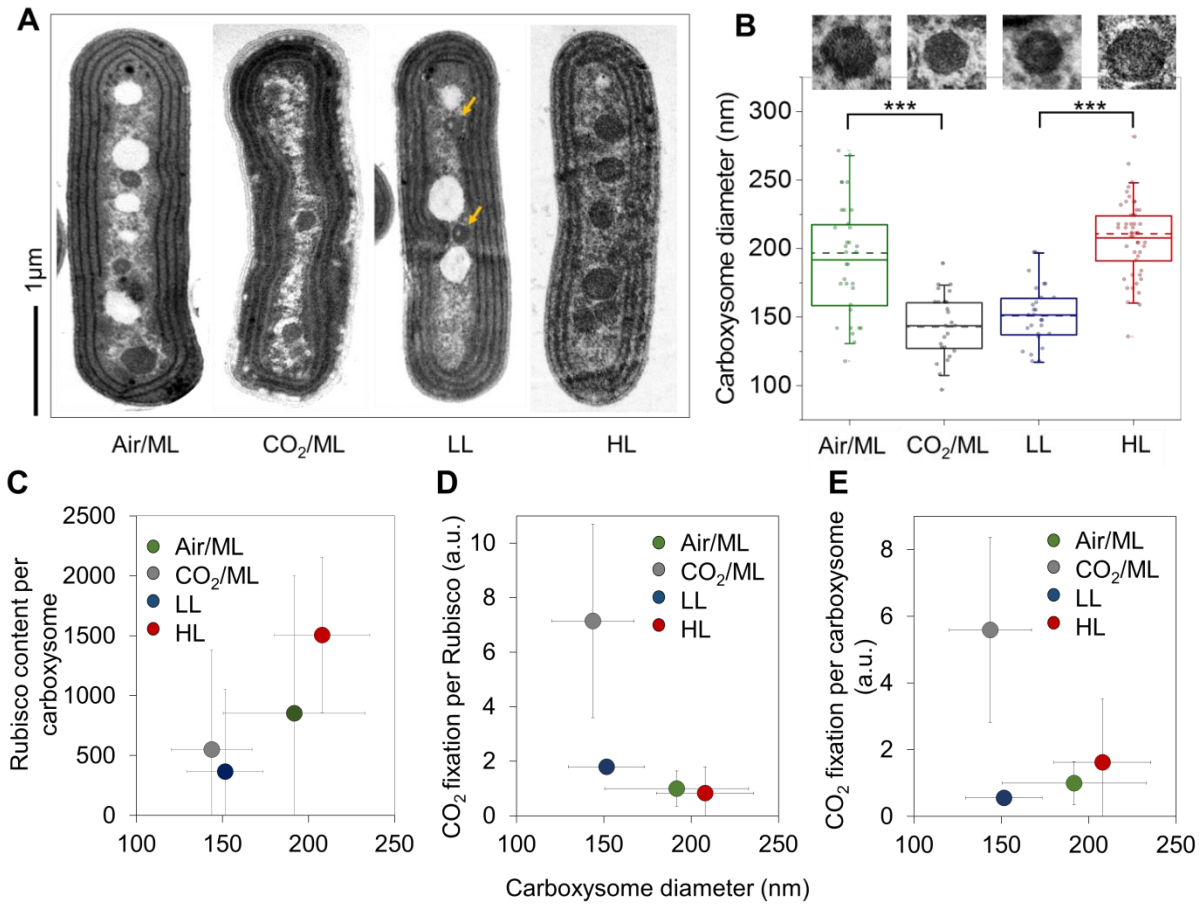


Figure 4-13. Variations of the carboxysome size and carbon fixation under Air/ML, CO_2, LL and HL. A. Thin-section electron microscopy images showing individual carboxysomes in the Syn7942 WT cells under Air/ML, CO_2, LL and HL treatments (Scale bar = 1 μm). See more images in **Figure 4-14**. Yellow arrows indicate the carboxysomes with spaces of low protein density under LL. **B.** Changes in the carboxysome diameter under Air/ML, CO_2, LL and HL measured from EM (*n* = 33, 25, 27 and 51, respectively), with representative carboxysome images above. Differences in the carboxysome diameter are significant between CO_2 and air ($p = 1.92 \times 10^{-14}$) and between LL and HL ($p = 8.29 \times 10^{-7}$), indicated as ***. Correlation between the carboxysome size and the Rubisco content per carboxysome (as **C**), CO_2 fixation per carboxysome (as **D**) and CO_2 fixation per Rubisco (as **E**) generated under Air/ML, CO_2, LL and HL. Carboxysome diameters and CO_2 fixation were present as average ± SD, whereas the carboxysome total protein content and Rubisco content were shown as Peak value ± HWHM.

Table 4-4. Properties of an average carboxysome in Syn7942 under Air/ML, CO_2/ML, LL and HL conditions. Data estimated from previous model (Long et al., 2011) are highlighted in grey.

Average CB properties	Air/ML	CO_2/ML	LL	HL
Carbon fixation rate per Rubisco (a.u.)	1.00	7.15	1.80	0.83
Coverage of CcmK3-K4 +CcmL on surface (%)	19.2	34.1	31.0	16.3
Ratio of CcmM: Rubisco	0.843	0.851	1.316	0.780
Ratio of CcaA: Rubisco	0.101	0.235	0.177	0.081
Copies of Rubiscos	853	550	367	1507
CB diameter (nm)	192	144	151	208
CB surface area (10^4 nm^2)	8.82	4.96	5.46	10.36
Internal CB volume (10^6 nm^3)	1.88	0.75	0.87	2.42
Surface area/volume ratio	0.039	0.052	0.050	0.036
Rubisco density in lumen (10^{-4} copy/nm^3)	9.08	11.80	6.75	9.95
Ratio of CcmK4:CcmK3	3.8	3.6	4.1	3.9

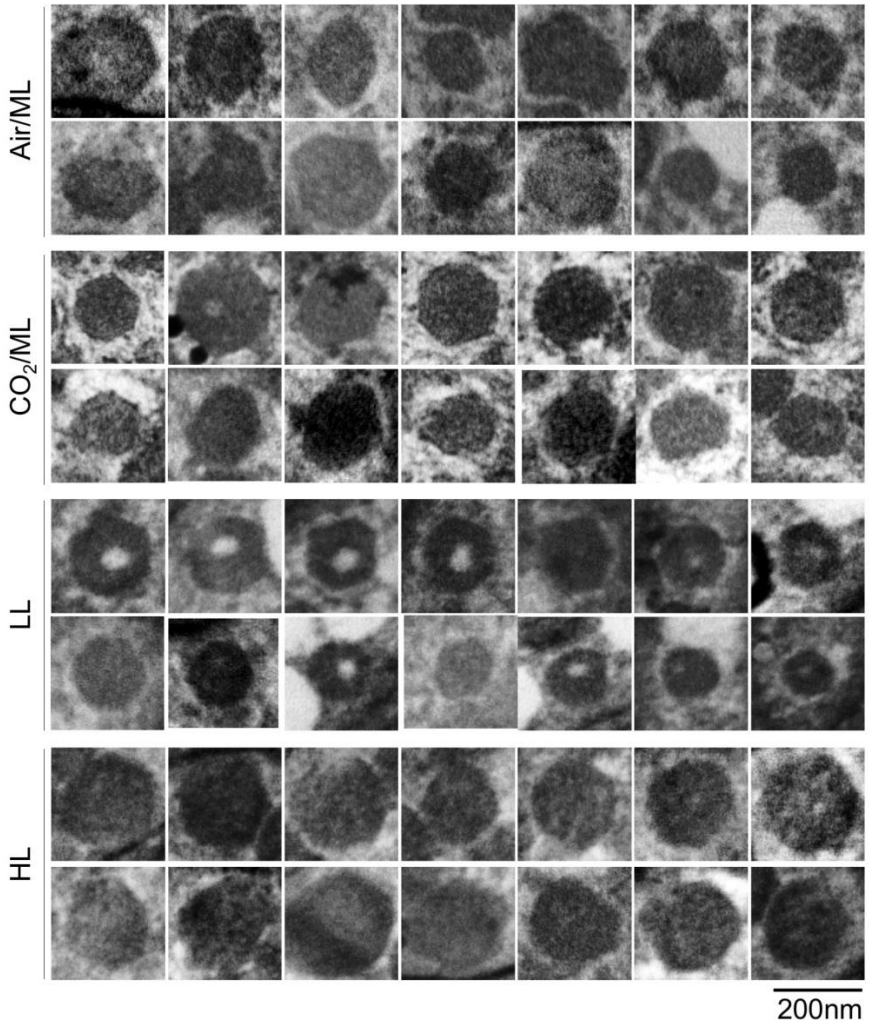


Figure 4-14. Thin-section electron microscopic images of Wild-Type Syn7942 cells under Air/ML, CO_2/ML, LL and HL. The size of carboxysomes in cells grown under different environmental conditions varies. Low density regions were often found in the lumen of carboxysomes synthesized under LL (59%, 16 out of 27 β-carboxysomes), compared with 9% for Air/ML (3 out of 33), 12% for CO_2/ML (3 out of 25) and 8% for HL (4 out of 51). Scale Bar = 200 nm.

4.2.4 The patterns of spatial localisation and diffusion of β-carboxysomes in live cells change dynamically depending upon light intensity during growth

The patterns of localisation of β-carboxysomes within the cyanobacterial cells appear to be crucial for carboxysome biogenesis and metabolic function (Savage et al., 2010; Sun et al., 2016). We measured the organisational dynamics of β-carboxysomes with distinct diameters in Syn7942 under different light intensities, using time-lapse confocal fluorescence imaging on the RbcL-YFP Syn7942 strain. Previous studies have shown that tagging of RbcL with fluorescent proteins does not obstruct β-carboxysome assembly and function in Syn7942 (Cameron et al., 2013; Chen et al., 2013; Savage et al., 2010; Sun et al., 2016). During time-lapse confocal imaging, we applied illumination with the same light intensity on the cell samples as that for cell growth to maintain cell physiology. We find that the diffusion of individual β-carboxysomes within cyanobacterial cells is non-Brownian (**Figure 4-15A**), indicative of intracellular restrictions that mediate carboxysome positioning. Carboxysomes under HL display larger diffusive regions than those under LL. The mean square displacement (MSD) of tracked carboxysomes increased with the rise of light intensity (**Figure 4-15B**), as did the mean microscopic diffusion coefficient of individual carboxysomes (**Figure 4-15C**): an average diffusion coefficient of $2.76 \pm 2.83 \times 10^{-5}$ $\mu m^2 \cdot sec^{-1}$ for HL (mean ± SD, n = 105), $1.48 \pm 1.03 \times 10^{-5}$ $\mu m^2 \cdot sec^{-1}$ for ML (n = 84), and $0.28 \pm 0.19 \times 10^{-5}$ $\mu m^2 \cdot sec^{-1}$ for LL (n = 336). Notably, carboxysomes with a larger diameter (**Figure 4-13**) generated under HL present a higher diffusion coefficient compares with carboxysomes with relatively smaller size under ML and LL. However, there is no apparent correlation between the diffusion coefficient of carboxysomes and their size in the same light conditions (**Figure 4-16**). These results indicate the mechanisms that dominate the spatial organisation and diffusion of carboxysomes in Syn7942, for example, the previously proposed interactions with cytoskeletal proteins (Savage et al., 2010) and ParA-mediated chromosome segregation (Jain et al., 2012), are also dependent sensitively on light intensity during cell growth.

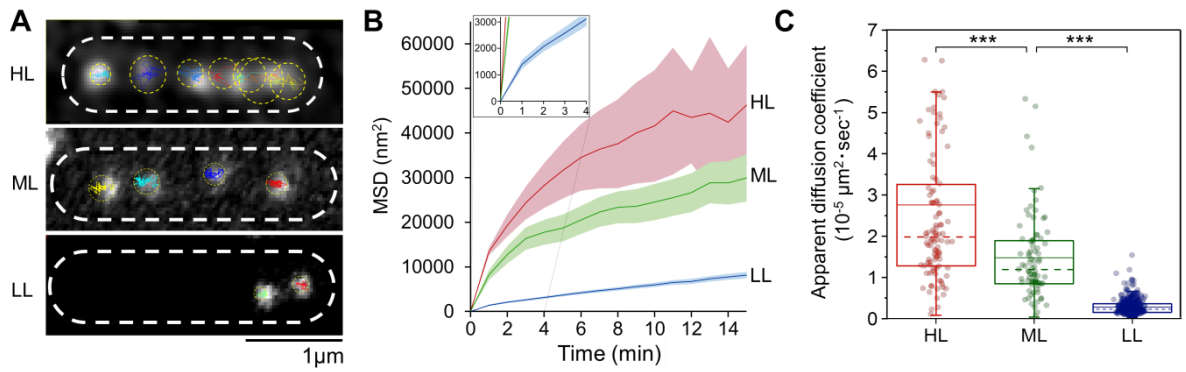


Figure 4-15. Spatial localisation and diffusion dynamics of carboxysomes in Syn7942 cells are dependent on light intensity. A. Tracking of carboxysome diffusion in cells grown under HL, ML and LL. Coloured lines indicate the diffusion trajectories of each carboxysomes and circles represent the diffusion areas of each carboxysomes over 60 mins. **B.** Non-linear MSD (Mean Square Displacement) profiles suggest the constrained diffusion of carboxysomes in Syn7942 cells grown under HL, ML and LL. Inset, zoom-in view of the MSD profile under LL. **C.** Diffusion coefficient of carboxysomes *in vivo* decreases significantly when the light intensity reduces: 2.76 ± 2.83 x 10^{-5} $\mu m^2 \cdot sec^{-1}$ for HL (mean ± SD, n = 105), 1.48 ± 1.03 x 10^{-5} $\mu m^2 \cdot sec^{-1}$ for ML (n = 84), and 0.28 ± 0.19 x 10^{-5} $\mu m^2 \cdot sec^{-1}$ for LL (n = 336). (p = 3.05 x 10^{-5} for comparison between HL and ML, p = 2.77 x 10^{-5} for comparison between ML and LL, two-tailed student t-test).

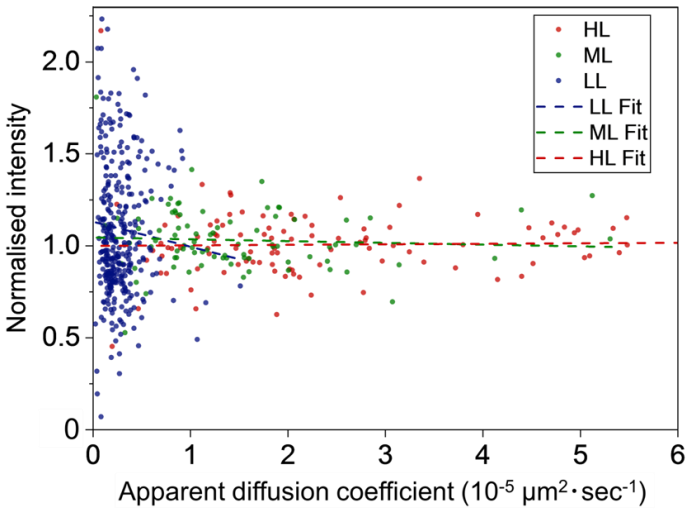


Figure 4-16. Changes in the diffusion coefficient of carboxysomes in Syn7942 cells under HL, ML and LL are not dependent on the carboxysome size. The correlation between diffusion coefficient and RbcL-YFP fluorescence intensities per carboxysome (representing the carboxysome size) under HL, ML and LL could not be well linearly fitted (R^2 < 0.005 for HL, ML and LL).

4.3 Discussion

Precise quantification of the protein stoichiometry and organisational regulation of carboxysomes provides insight into their assembly principles, structure and function. In this work, we functionally fused fluorescent protein tags to the building blocks in β-carboxysomes and exploited single-molecule fluorescence microscopy to count the actual protein stoichiometry of β-carboxysomes in Syn7942 cells, at the single-organelle level. We characterised the stoichiometric flexibility of carboxysome proteins within the polyhedral structure towards environmental variations. Variability of the protein stoichiometry and size of carboxysomes likely provide the structural foundation for the physiological regulation of carboxysome biogenesis and carbon fixation activity. Given the shared structural features of carboxysomes and other BMCs, we believe that this work opens up new opportunities to quantitatively evaluate protein abundance and decipher the formation of all BMC organelles, in both native forms and synthetic variants.

Despite a great deal of prior efforts on understanding carboxysome structure and function, the relative stoichiometry of functional carboxysome components in their native cell environment − key information required for reconstituting active carboxysome structures in synthetic biology (Fang et al., 2018), was still unclear. The major challenges have been the unspecificity of immunoblot and mass spectrometry given the homology of carboxysome proteins and the lack of effective purification of entire carboxysomes from host cells as well as the heterogeneity of carboxysome structures (Long et al., 2005). Recent work in our lab reported the isolation of β-carboxysomes from Syn7942 and the structural and mechanical exploration of the organelles (Faulkner et al., 2017). Interestingly, some components, i.e. CcmO, CcmN, CcmP and RbcX, were not detected by mass spectrometry in the isolated carboxysomes, likely due to their low content and potential destruction of the carboxysome structure during isolation. Here, as demonstrated, fluorescence tagging and Slimfield and confocal imaging enable single-organelle analysis of the protein stoichiometry of 7 β-carboxysomal proteins and their regulation in their native context and extends analyses of the assembly and action of carboxysomes. Microscopic imaging on fluorescently-tagged β-carboxysomes has been used to reveal

their patterns of cellular localisation, biogenesis pathways and light-dependent regulation in Syn7942 (Cameron et al., 2013; Chen et al., 2013; Savage et al., 2010; Sun et al., 2016). We validate that YFP tagging to most of the structural components does not impede the formation of functional carboxysome structures, suggesting the functional relevance of the determined protein stoichiometry in the carboxysome. Exceptionally, fluorescence tagging on BMC-T proteins CcmP and CcmO does not show normal carboxysome assembly and localisation compared to other YFP-tagged strains (**Figure 4-17**). In this work, therefore, we did not include estimation of the protein abundance of CcmP and CcmO, as well as RbcL and CcmK2 that cannot be fully tagged with YFP, using Slimfield and confocal imaging.

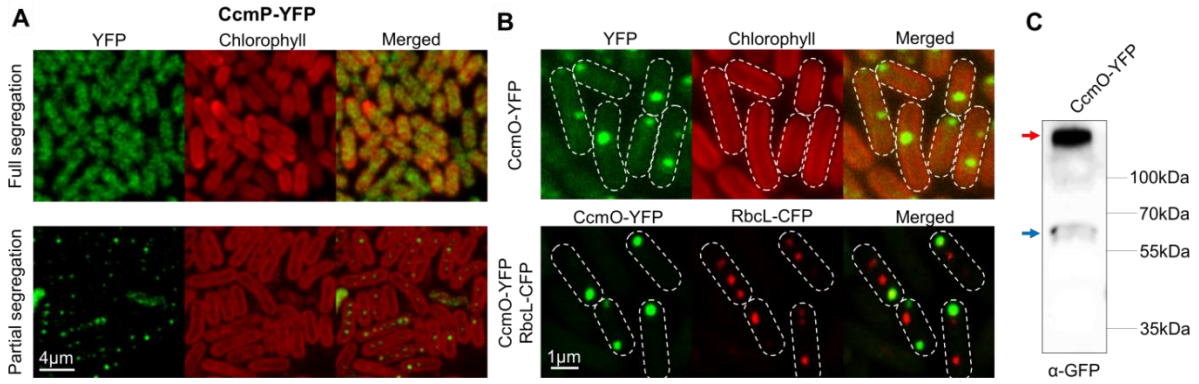


Figure 4-17. CcmP-YFP and CcmO-YFP Syn7942 cells. A. Fully-segregated CcmP-YFP cells show abnormal carboxysome distribution and assembly, whereas partially-segregated cells present typical carboxysome distribution. **B.** Only 1-2 fluorescence foci are seen in the fully-segregated CcmO-YFP cells fusion. Confocal imaging of the double-labelled CcmO-YFP::RbcL-CFP strain reveals that CcmO-YFP and RbcL-CFP are not co-localised in the cell. **C.** Immunoblot of soluble fractions of the CcmO-YFP strain using an anti-fluorescent protein antibody, showing the presence of large aggregates of CcmO that failed to run into the gel, marked by red arrow. Minor fractions of CcmO that run into the gel are marked by blue arrow. For all western results, at least 3 independent experiments were done for statistical analysis.

Numerous studies have described the regulation of carboxysome protein expression at the transcriptional level (McGinn et al., 2003; Schwarz et al., 2011; Woodger et al., 2003). Counting protein abundance of β-carboxysomes at different cell growth conditions enables direct characterisation of the stoichiometric plasticity of carboxysome building components in the cells

grown under not only the same environmental condition but also various conditions (**Figure 4-18**). It appears that the regulation of protein abundance within the carboxysome correlate with the genetic locations and organisation of carboxysome genes, indicative of stoichiometric adjustment of carboxysome building components. Our observations further elucidate the size variation of β-carboxysomes in Syn7942 cells grown under distinct environmental conditions (**Figure 4-18**) and adjustable carbon fixation capacities of carboxysomes that may be closely linked to the protein organisation and size of carboxysomes. Variations in the diameter of intact carboxysomes, ranging from 90 to 600 nm, have been also shown in previous studies not only in single species but also among distinct species, suggesting the adaptation strategies exploited by cyanobacteria for regulating their CO_2-fixing machines to survive in diverse niches (Iancu et al., 2007; Liberton et al., 2011; Price and Badger, 1991; Shively et al., 1973). Moreover, the spatial positioning and mobility of β-carboxysomes in live cells appear to be independent of carboxysome diameter but show a strong dependence to light intensity, suggesting the potential light-dependent mechanisms that mediate carboxysome location and diffusion. Previous studies have illustrated that the cytoskeletal proteins ParA and MreB are mediators of carboxysome positioning in Syn7942 and equal segregation between daughter cells (Savage et al., 2010). In addition, carboxysome spacing and partitioning were suggested to be driven by different possible mechanisms, including ParA-mediated chromosome segregation (Jain et al., 2012) via filament-pull model (Ringgaard et al., 2009) or a diffusion-ratchet model (Vecchiarelli et al., 2013). Altogether, the organisational flexibility of β-carboxysomes, including modulable protein stoichiometry, diameter and mobility, may represent strategies for modulating shell permeability and enzyme encapsulation and ensuring structural and functional adaptations dependent on the local cellular environment.

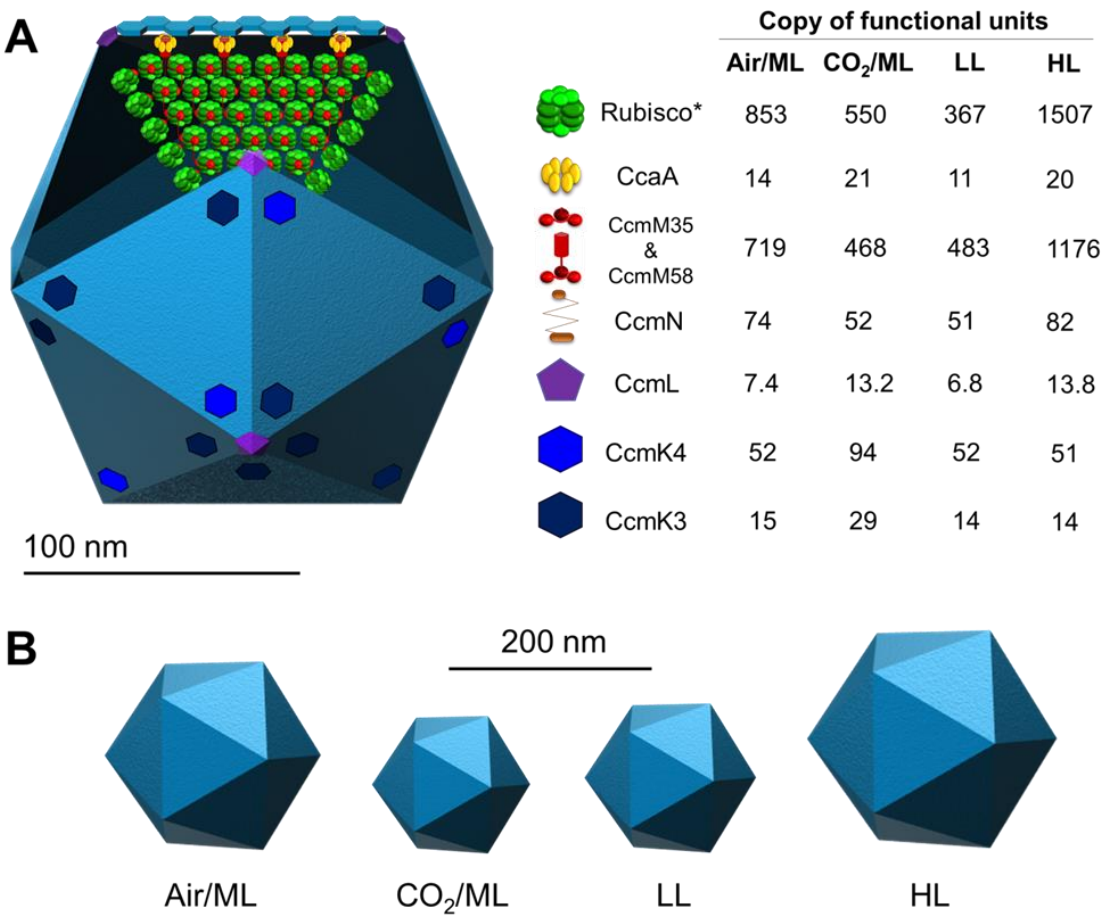


Figure 4-18. Model of the β-carboxysome structure and protein stoichiometry. A. Diagram of an icosahedral carboxysome structure and organisation of building components. The stoichiometry of each building component within the carboxysome and its variations in response to changes in CO_2 and light intensity are shown on the right. CcmO and ccmP are not included in the model as they were not addressed in this work. Majority of the shell in light blue colour represent the major shell protein CcmK2. **B.** The carboxysome diameter is variable in response to changes in CO_2 and light intensity. *Rubisco contents are estimated from RbcS quantify and based on a L_8S_8 formation.

The estimated number of CcmL pentamers per carboxysome could be less than 12, demonstrating explicitly that it is not a prerequisite for CcmL pentamers to occupy all 12 vertices of the shell to ensure complete formation of functional carboxysomes. This hypothesis has been validated by previous observations that BMC shells in the absence of pentamers have no significant morphological changes (Cai et al., 2009; Hagen et al., 2018; Lassila et al., 2014). These "non-complete capped" forms appear to be prevalent among the resultant carboxysomes under Air/ML and LL (**Figure 4-8C**), unlike the procarboxysomes (Cameron et al., 2013) or "immature" carboxysomes which are incapable

of establishing an oxidative microenvironment for cargo enzymes (Chen et al., 2013). It suggests that carboxysomes could possess a flexible molecular architecture, which resonates with the observation of structural "breathing" of virus capsids which has been reported to be key to cope with temperature change (Li et al., 1994; Roivainen et al., 1993). Carboxysomes, though structurally resembling virus capsids, have been shown to be mechanically softer than the P22 virus capsid by a factor of ~10, suggesting greater flexibility of protein-protein interactions within the carboxysome structure (Faulkner et al., 2017). The capping flexibility of pentamers may represent the dynamic nature of shell assembly probably in the second scale and tunable protein-protein interactions in the shell, as characterised recently (Sutter et al., 2016).

Based on the RbcS abundance per carboxysome, we estimated that there are approximately 853, 550, 367, and 1507 copies of Rubisco per β-carboxysome under Air/ML, CO_2/ML, LL, and HL, respectively (**Figure 4-18**, **Table 4-2**). Even the lowest Rubisco abundance per β-carboxysome (average diameter of 151 nm) under LL is still greater than the Rubisco abundance per α-carboxysome (average diameter of 123 nm) (Iancu et al., 2007) by a factor of 1.6. This finding confirms the different interior organisation of the two classes of carboxysomes; densely packed of Rubisco that forms paracrystalline arrays inside the β-carboxysome (Faulkner et al., 2017) and random packing of Rubisco in the α-carboxysome (Iancu et al., 2007; Iancu et al., 2010). The different interior structures of α- and β-carboxysomes may be ascribed to their distinct biogenesis pathways: biogenesis of β-carboxysomes is initiated from the nucleation of Rubisco and CcmM35 and then the shell encapsulation (Cameron et al., 2013); whereas α-carboxysome assembly appears to start from shell formation (Menon et al., 2008) or a simultaneous shell-interior assembly (Iancu et al., 2010).

While the abundance of most of the structural components varies in the manners that are likely related to the locations of corresponding genes in the genome, the ratio of CcmK4 and CcmK3 is relatively unaffected (ranging from 3.6 to 4.1, **Table 4-4**) under the tested growth conditions, implying their spatial colocalisation within the carboxysome shell. The *ccmK3* and *ccmK4* genes are located in the

same operon that is distant from the *ccm* operon and they may have different expression regulation compared with other carboxysome components (Rae et al., 2012; Sommer et al., 2017). The balanced expression and structural cooperation of CcmK3 and CcmK4 may be crucial for the fine-tuning of carboxysome activity and permeability towards environmental stress.

Rational design, construction and modulation of bioinspired functional materials are the major challenges in synthetic biology and protein engineering. Given their self-assembly, modularity and high efficiency in enhancing carbon fixation, carboxysomes have attracted tremendous interest to engineering this CO_2-fixing organelle into other organisms, for example C_3 plants, with the intent of increasing photosynthetic efficiency and crop production (Lin et al., 2014a; Lin et al., 2014b; Long et al., 2018; Occhialini et al., 2016). Recently, we have reported the engineering of functional β-carboxysome structures in *E. coli* – a step towards constructing functional β-carboxysomes in eukaryotic organisms (Fang et al., 2018). Our present study, by evaluating the actual protein stoichiometry and structural variability of native β-carboxysomes, sheds light on the molecular basis underlying the assembly, formation and regulation of functional carboxysomes. It will empower bioengineering to construct BMC-based nano-bioreactors and scaffolds, with functional and tunable compositions and architectures, for metabolic reprogramming and targeted synthetic molecular delivery. A deeper understanding of carboxysome structure and the developed imaging techniques will be broadly extended to other BMCs and macromolecular systems.

Chapter 5

Biosynthesis and organisation of cyanobacterial carbon fixation machinery in diurnal light: a glance into the regulation of circadian clock

Chapter 5

5.1 Introduction

5.1.1 The circadian clock in cyanobacteria

Like many other organisms that experienced day-night shifting in nature, cyanobacteria have evolved internal acclimation mechanisms to aid in the orchestration of activities specific to light-dark period. For a large proportion of Syn7942 genes, rhythmic changes at transcriptional level have been detected during light-dark daily cycles (Golden, 2003). The different expression profiles among genes over the whole genome can be categorised into three groups: the Class I genes which have the highest level of expression during dusk before entering darkness, the Class II genes which have the highest expression during dawn before light illumination, and the arrhythmic genes with non-rhythmic expression profiles (Ito et al., 2009). The rhythmic expression profiles of Class I and II genes can be sustained even after removal of cyclic input, for instance, after switching the cells to constant light condition (hereafter CL) (Ditty et al., 2005). The internal biological clock system which was believed to be eukaryote-specific previously has been identified in Syn7942, extending the understanding of circadian control in bacterial phyla (Ishiura et al., 1998). The cyanobacterial circadian clock consists of three genes named *kaiA/B/C* (the word "Kai" in Japanese means "cycle"). Homologs of *Kai* were then discovered in diverse prokaryotic organisms (Loza-Correa et al., 2010). The self-maintained core clock complex is composed of three proteins, the KaiA/B/C, as illustrated in **Figure 5-1**. In general, the KaiC hexamers go through periodically phosphorylation and dephosphorylation in a 24-hour period fashion continuously through interaction with KaiA and KaiB (Nishiwaki et al., 2004). KaiA binds to the A-loop of KaiC and stimulates KaiC autophosphorylation during the daytime, whereas during the night time, KaiB competitively binds to KaiC and releases KaiA into inactive forms, leading to phosphorylation of KaiC (Goda et al., 2014). Together, such post-translational oscillator (PTO) forms the master clock that regulates gene expressions (Swan et al., 2018).

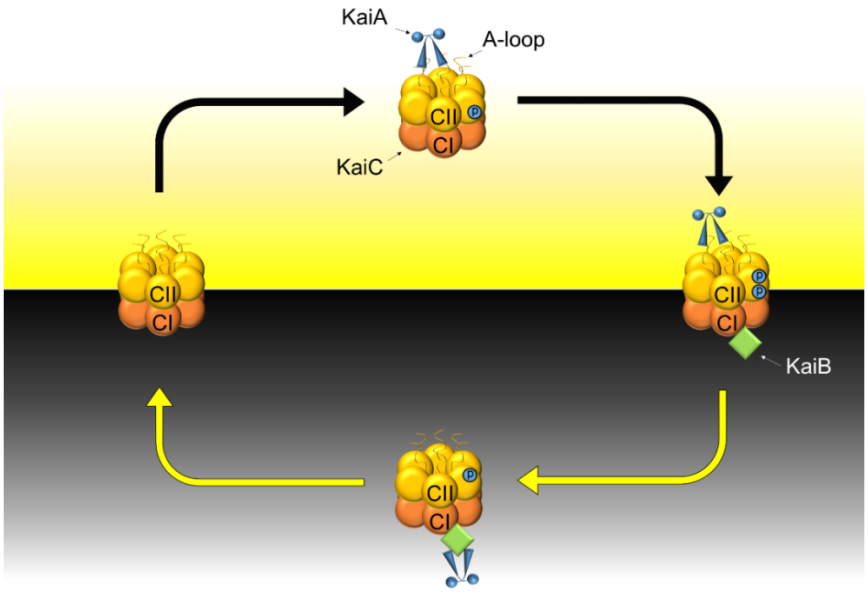


Figure 5-1. Mechanism of the Kai circadian clock oscillator. Cyanobacterial circadian clock core oscillator is consist of three protein, KaiA/B/C. During daytime, the KaiC hexamers are phosphorylated by binding with KaiA through A-loop. During night time, KaiC hexamers are dephosphorylated by binding with KaiB and releasing of KaiA from A-loop, associated with KaiB. The periodical phosphorylation/ dephosphorylation can be self-maintained in a 24-hour fashion.

Strikingly, such biological clock through the rhythmic phosphorylation of KaiC could be reproduced *in vitro* solely with the core components KaiA/B/C and ATP. Such fact indicates the standalone and robust persistency for pace keeping (Nakajima et al., 2005). In cells, the circadian clock system function through three components, the core oscillator *kai*, the input mechanism that transmits environmental stimuli to the core oscillator which synchronise and reset of the clock, and the output mechanism that signals diverse cellular activities (Taniguchi et al., 2010). The structural basis and input/output systems of the circadian clock in Syn7942 are well reviewed (Cohen et al., 2015; Johnson et al., 2017). A variety of environmental signals could feedback to the core clock through redox-sensitive components such as CikA and LdpA (Ivleva et al., 2005; Ivleva et al., 2006). Besides, KaiA and KaiC proteins in the core clock are also capable of sensing the cofactors in the photosynthetic activity directly without the previously described redox-sensing input pathways (Rust et al., 2011; Wood et al., 2010). Through such combined sensing mechanisms, the core clock can be effectively synchronised to the environment.

Interestingly, CikA was later found to be involved in the circadian output system (Gutu and O'Shea, 2013). The output histidine kinase proteins SasA/RpaA and CikA sense the phosphorylation status of the core oscillator directly through protein-protein interactions during the day and night periods respectively. In day period, SasA binding to KaiC promotes the phosphorylation of RpaA and leads to the accumulation of phosphorylated RpaA throughout the day. The phosphorylated RpaA then promotes the translation of Class I genes to reach maximal transcription during dusk and inhibits the translation of Class II genes to reach peak transcription during dawn (Iwasaki et al., 2000; Takai et al., 2006). During the night period, oxidative redox state leads to the aggregation of KaiA/C to the membrane where CikA and KaiA interact with the quinones, preventing the SasA/RpaA transmission pathway, while a RpaA co-modulator RpaB takes charge for inhibition of Class I gene (Espinosa et al., 2015). Periodically aggregation of core complexes are found forming at the cell pole to interact with quinones during dark period (Cohen et al., 2014). **Figure 5-2** shows the diagram of regulatory mechanism described above.

Under diurnal condition, cells benefit from the circadian regulation and show improved overall fitness, directly represented by faster growth (Woelfle et al., 2004). It is commonly acknowledged that such benefits come from the forecasted anticipation of incoming darkness or light and prepare the cell beforehand (i.e., protein synthesis, reorganisation of complexes and pathways), therefore grant cells advantages than those who merely answer to the environmental stimuli (Cohen and Golden, 2015).

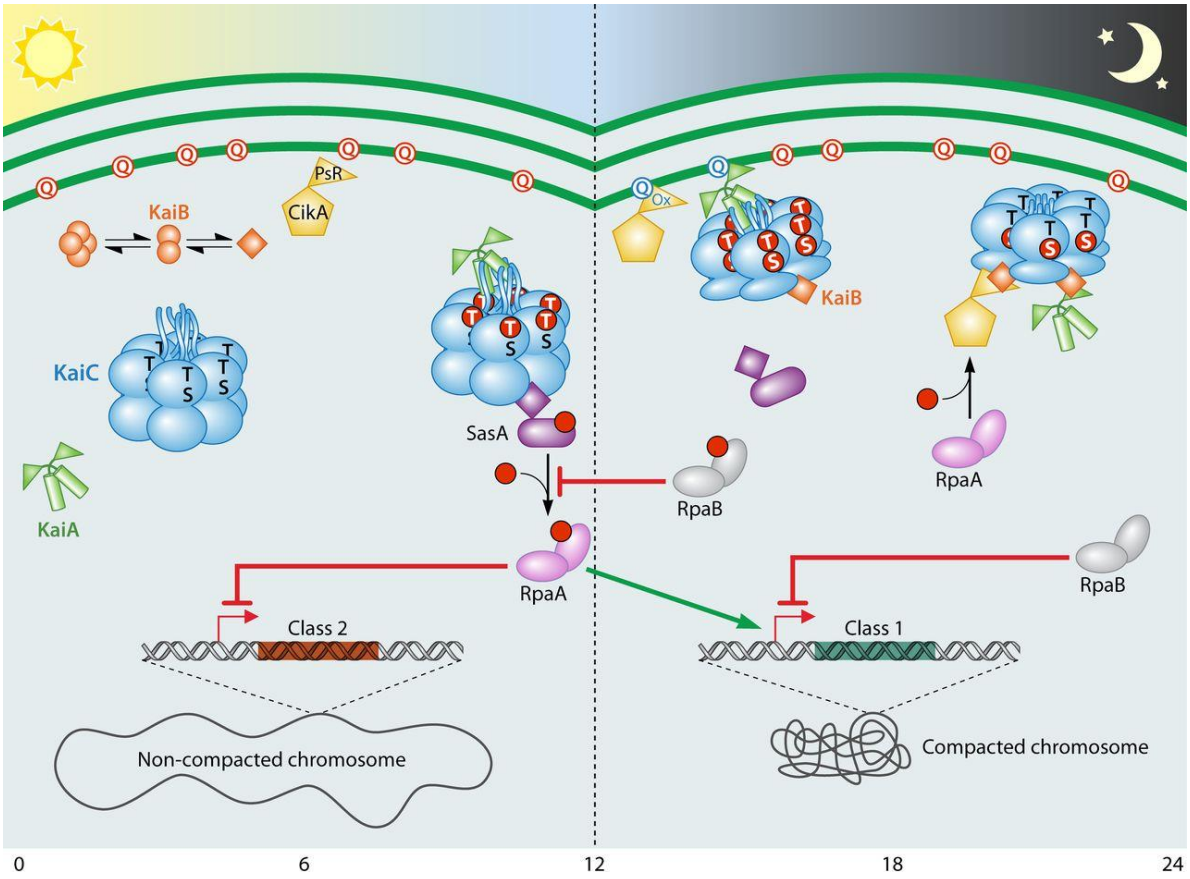


Figure 5-2. Model of circadian control mechanism during light-dark cycles within the cell. (Cohen and Golden, 2015). Diagrams shown on the left and right of the dashed line indicate the respective model during day/night conditions. In the day period, cytosolic KaiA promotes KaiC phosphorylation (illustrated in red sphere over the CII domain). SasA interacts with KaiC, promoting phosphorous transfer to RpaA. Accumulation of phosphorylated RpaA reach peaking quantity at the end of light period, progressively suppressing the expression of class II genes (dawn-peaking) and activates class I genes (dusk-peaking). In the night period, KaiA/C interact with CikA and oxidative quinones over the cell poles. KaiB binds to KaiC and releases KaiA, promoting KaiC dephosphorylation, illustrated as the reducing content of red sphere over KaiC. The formed KaiB/C complex promotes RpaA phosphorylation through CikA. RpaB inhibits the expression of class I gene in dark period, and crosstalk with RpaA further inhibits the phosphorylation of RpaA during the night. Redox state of quinones in the thylakoid membranes are shown in red/blue colours indicating reduced and oxidised states respectively.

5.1.2 Role of Rubisco coupling with the rhythmic diurnal cycles

The fitness improvement provided by circadian control has also been investigated in the photosynthetic systems. In the model plant *Arabidopsis thaliana*, circadian regulation results in the increased chlorophyll concentration in leaves, enhanced photosynthesis and biomass production, as

well as the increase in CO_2 assimilation under a well-synchronised clock that matches to the corresponding timing of the dark-light conditions (Dodd et al., 2005). Among all photosynthetic organisms, Photosynthesis cope with the diurnal cycle through a class II-type regulation, following a similar strategy that ready the photosynthetic system before sun rises. On the contrary, distinctive strategies have been found in carbon fixation processes among different species. For example, In some plants, a circadian regulated mechanism called Crassulacean acid metabolism (CAM) is adopted to diminish the competition of O_2 generated through photosynthesis during the daytime and to enhance CO_2 fixation efficiency via the production and storage of four-carbon malic acid at night (Ting, 1985). CO_2 fixation is divided in two steps: the primary fixation fulfilled by phosphoenolpyruvate carboxylase (PEPC) that stores inorganic carbon in the form of malic acid during the night with stomata opened for Ci uptake, and the secondary fixation by Rubisco using CO_2 pre-collected in malate acid and allows stomata remain closed during the day (Michelet et al., 2013). Such night-based Ci uptake and the primary fixation reduce the water losses, thus enhancing the survivability of plants in arid environment (Eller and Ferrari, 1997). Functional circadian clock in CAM plant can enhance and prolong the primary fixation towards malate in dark through the circadian clock-controlled protein Phosphoenolpyruvate Carboxylase kinase (PPCK). Silencing of PPCK halves the carbon fixation by PPC and lead to arrhythmia of the central circadian clock, suggesting strong cross-talk between PPC carbon fixation and the circadian clock in CAM plants (Boxall et al., 2017). Moreover, the CO_2 fixation activity by Rubisco has shown to be under the circadian control inferred from the rhythmic CO_2 uptake patterns in continuous light in CAM plant (Wyka and Luttge, 2003). Beside CAM plants, the activity of carbon fixation can be orchestrated by circadian over the daily cycle through re-distribution of Rubisco inside the chloroplasts while maintaining a relatively unaltered Rubisco content in a chloroplast-containing marine algae (Nassoury et al., 2001).

However, regarding the carbon fixation activity of cyanobacteria under diurnal regulation, far lesser reports could be found, particularly ones further discuss the circadian involvement. In Syn7942, even though several studies using transcriptomic analysis (Ito et al., 2009; Vijayan et al., 2011; Vijayan et

al., 2009) categorised carboxysomal genes as circadian-regulated Class II genes, no further evidence have been provided regarding the variations of abundance for major carboxysomal protein such as RbcL, RbcS and CcmK during the diurnal cycles. Interestingly the content of CcaA seems remains unaltered indicated from high throughput Mass Spectrometry in Syn7942 (Guerreiro et al., 2014) and the cyanobacterium Cyanothece ATCC511 (Aryal et al., 2011). Meanwhile transcriptomic data has suggest rhythmic *ccaA* transcriptions. The discrepancies found between transcriptional and protein levels suggest unknown post-transcriptional regulation for carboxysomal gene under diurnal conditions. The difference of CO_2 fixation activity of cells between the diurnal condition and CL condition, as well as the involvement of circadian-originated transcriptional regulation over motabolic processes by carboxysomes remain largely undocumented.

In this chapter, we documented the *in vivo* localisation, CO_2 fixation activity of carboxysomes. We further quantified the content of Rubisco under both diurnal dark and light conditions to know how carboxysomes are generated in cyclic changed environment. Moreover, through characterisation of a circadian nullified strain, we gathered useful information regarding the involvement of circadian control in carboxysome biogenesis in Syn7942.

5.2 Results

To ensure the proper diurnal light and dark treatments could be achieved on the platform we designed, pAM2195 (Mackey et al., 2007) that express both *luxAB* and *luxCDE* with promoter promoted by P*psbA*I which enable host strain produce bioluminescent autonomously was transformed into WT Syn7942. Rhythmic changes of luciferase bioluminescence were observed within the 22 h period (**Figure 5-3A**). The profiles of signal intensity exhibit the intensity peak at 12 h (**Figure 5-3B)**, consistent with previous results (Mackey et al., 2007). Furthermore, the growth curves of cells in liquid culture under CL were recorded (**Figure 5-3C**). The pAM2195 mutant exhibited WT levels of cell doubling time, with an acceptable 10% increase caused by the expression of luciferase under

strong promoter *psbAI*. Nevertheless, the rhythmic changes of luciferase indicate active circadian clock control on our established platform.

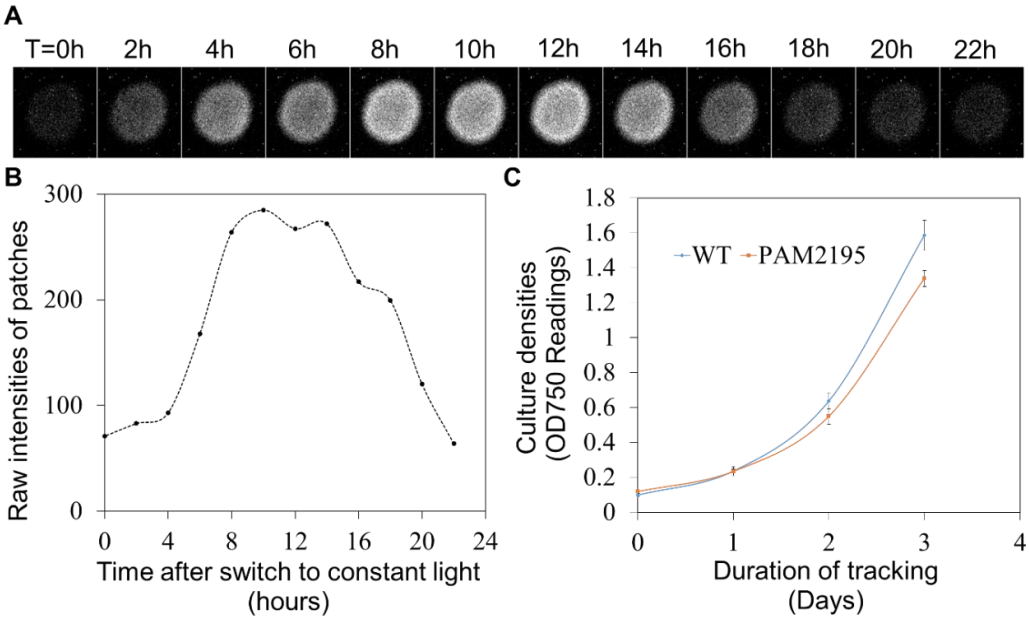


Figure 5-3. Confirmation of circadian control in pAM2195 luciferase reporter strains. A. Time-lapse imaging for patches of Syn7942 cells with luciferase reporter vector pAM2195 under CL after dark end after switched from circadian light treatment (12L12D). Two-hour interval time was used to capture the luciferase illumination. Analysis of the rhythmic brightness shows that the strongest signal intensity appears at the subjective light 12h and reduces throughout the subjective dark (remained in light), indicating the proper circadian control; **B.** Quantification of luciferase illumination brightness on images taken during time-lapse imaging; C. Growth curve of WT and pAM2195 strains indicate acceptable inhibition on cell growth caused by luciferase expression. The doubling times are 17.9 ± 0.7 and 20.0 ± 0.7 hours respectively (n = 4).

To determine the involvement of circadian in carboxysome biogenesis in Syn7942 cells, a circadian null strain was generated through deletion of the core oscillator gene KaiA (Paddock et al., 2013). Also, KaiA was also fluorescently tagged at the C-terminal. Subcellular localisation relative to carboxysomal proteins can be retrieved. Moreover, correlations of carboxysome localisation to particular states of circadian could be established inferred from localisation of Kai complexes as reported previously (Cohen et al., 2014). In this chapter, we chose RbcL to visualise carboxysomes due to the high levels of signal intensities as well as the well documented localisation profiles in

previous work (Niederhuber et al., 2017; Savage et al., 2010; Sun et al., 2016). The mutant strains were validated through PCR segregation screening (**Figure 5-4**).

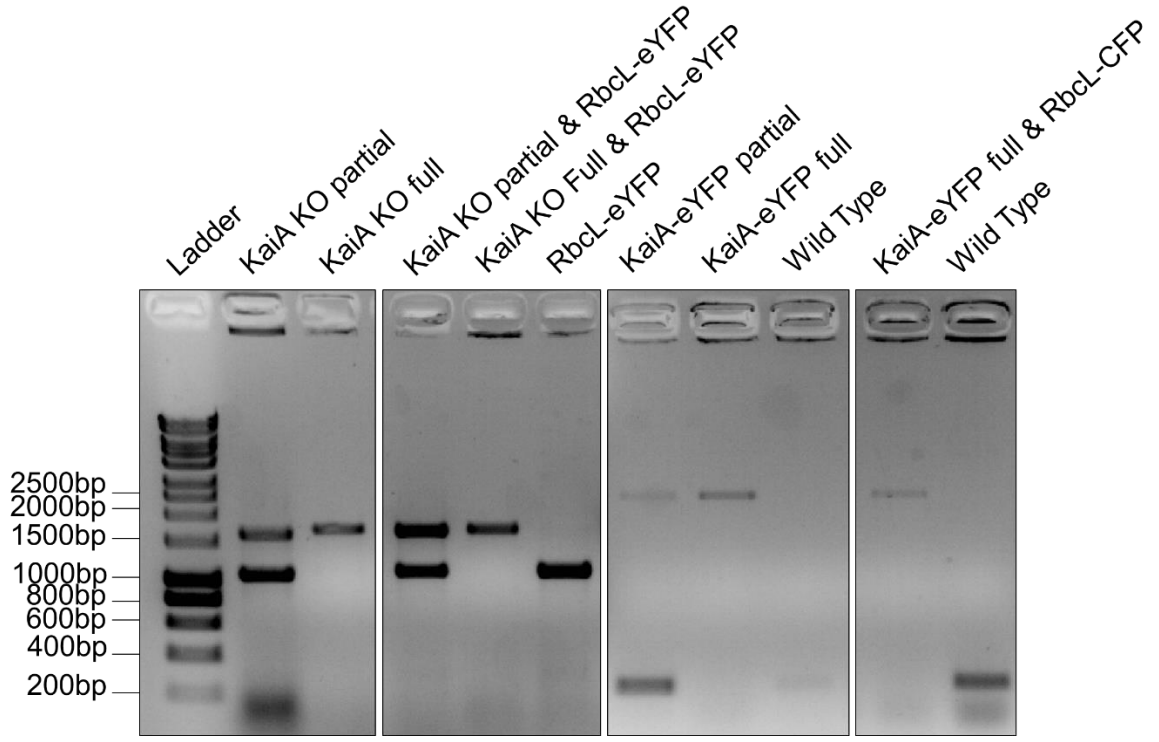


Figure 5-4. Segregation screening of KaiA mutants. For KaiA KO, the WT band size=1045 bp while KO band size= 1602 bp; For KaiA-eYFP, the WT band size=282 bp while YFP segregated band size=2408 bp. Only fully segregated strains were used for further examination. Knock-out indicated as KO.

5.2.1 The carboxysome localisation answering to the redox state of photosynthetic electron transport during diurnal conditions

It has been speculated that the positioning of carboxysomes within cells have crucial roles guaranteeing the even distribution of carboxysomes in daughter cells during cell division (Savage et al., 2010). In **Chapter 3**, we show drastic changing of carboxysome localisation along the short axis of the cell, preferring a more centralised positioning closer to cell axis in response to the reduced state of the PQ pool stimulated by DBMIB inhibitor. In nature, the PQ pool is periodically shifted from an oxidised state to reduced state during light and dark in a daily cycles (Schuurmans et al., 2014). We

wonder whether such stimulated rearrangement could occur naturally answering to altered state of PQ pool. Therefore, we recorded the localisation of carboxysomes across the diurnal cycles (12D12L) by confocal microscopy (**Figure 5-5A**). Eight-time points (D1H, D4H, D8H, D11H, L1H, L4H, L8H and L11H) were taken in total during a daily cycle. A more linear positioning along the short axis of the cell was observed during later hours in the dark (D8H and D11H) directly from the distribution curve in **Figure 5-5B**. The relative area further quantified the preferences of central line localisation under the distribution curve (**Figure 5-5D)**. We observed slightly centralised profiles from L11H to D4H and a significant and gradual increase in centralised distribution throughout the dark period from D8H to D11H. A swift recovery of less centralised distribution appears at L1H. The distribution profile remains at a similar level throughout the light period from L1H to L11H. The delayed repositioning phenomenon after 4 hours in dark corroborates the delayed repositioning after 4 hours of DBMIB treatment captured during the time-series imaging (**Chapter 3**, **Figure 3-17**). The re-organisation of carboxysome away from the cytosol peripheral to the centre of cells naturally occurs in the dark period with a minimum of 4 hours delay and returns near to the cytosol peripheral distribution status for respective light period within 1 hour.

Moreover, two-score system was established to analyse the distribution profile of carboxysomes along the long axis of the cell (**Figure 5-5C**) (described in **Chapter 2**, **Section 2.6.3**). The spatial distribution score evaluates the homogeneity of carboxysome distribution, where a higher score suggests specific localisation of carboxysomes along the long axis of the cell, shown as more distinguishable peaks with higher frequencies in the distribution curve, whereas random distribution results in a lower score, visualised as a more flattened distribution curve. The homogeneity of carboxysome distribution indicates the degree of control of carboxysome positioning within the cell. To evaluate the significance of score differences, we divided the cells imaged during each time point into three sub-group randomly, the scores were analysed for sub-group, and the systematic errors were calculated. For each time point, a minimum of 300 cells was analysed. The average system errors for different time points regarding the scores in **Figure 5-5E and F** are 0.011 and 0.014 respectively. Less spatial-like distribution was inferred from the lowing scores at the end of the light period at 0.16,

corresponding to L11H and the beginning of the incoming dark period at 0.10, corresponding to D1H (**Figure 5-5E**). At the same time, the longitudinal positioning control enhanced throughout the dark period, from D1H to D11H (0.10, 0.17, 0.20, 0.24). A sudden loosen of such control could be inferred when entering light period at L1H after D11H (0.24 to 0.19). Subsequentially, the control resumed gradually until L8H (0.19, 0.23 and 0.24), followed by a sudden loosening of the control again during L11H at 0.16.

The cell poles are suggested to be the birthplace of new carboxysomes (Cameron et al., 2013; Chen et al., 2013; Savage et al., 2010). Some essential cellular processes such as protein binding and recognition also mediated over pole regions (Laloux and Jacobs-Wagner, 2014). Likewise, the circadian core oscillator Kai was reported to be periodically aggregated to the cell poles during the dark period (Cohen et al., 2014). Therefore, we wonder whether there is any preference of carboxysome localisation along the long axis of cells in an assumption that newly generated carboxysomes are more likely to locate near the cell pole. However, we did not detect any significant changes (differences greater than 15% from the score) in the pole/central positioning (Polarity score = 0.21 to 0.27), but still during D11H an increased score at 0.27 that beyond systematic errors unmatched the trend line might suggest a slightly more “closer-to-pole” preference (**Figure 5-5F**).

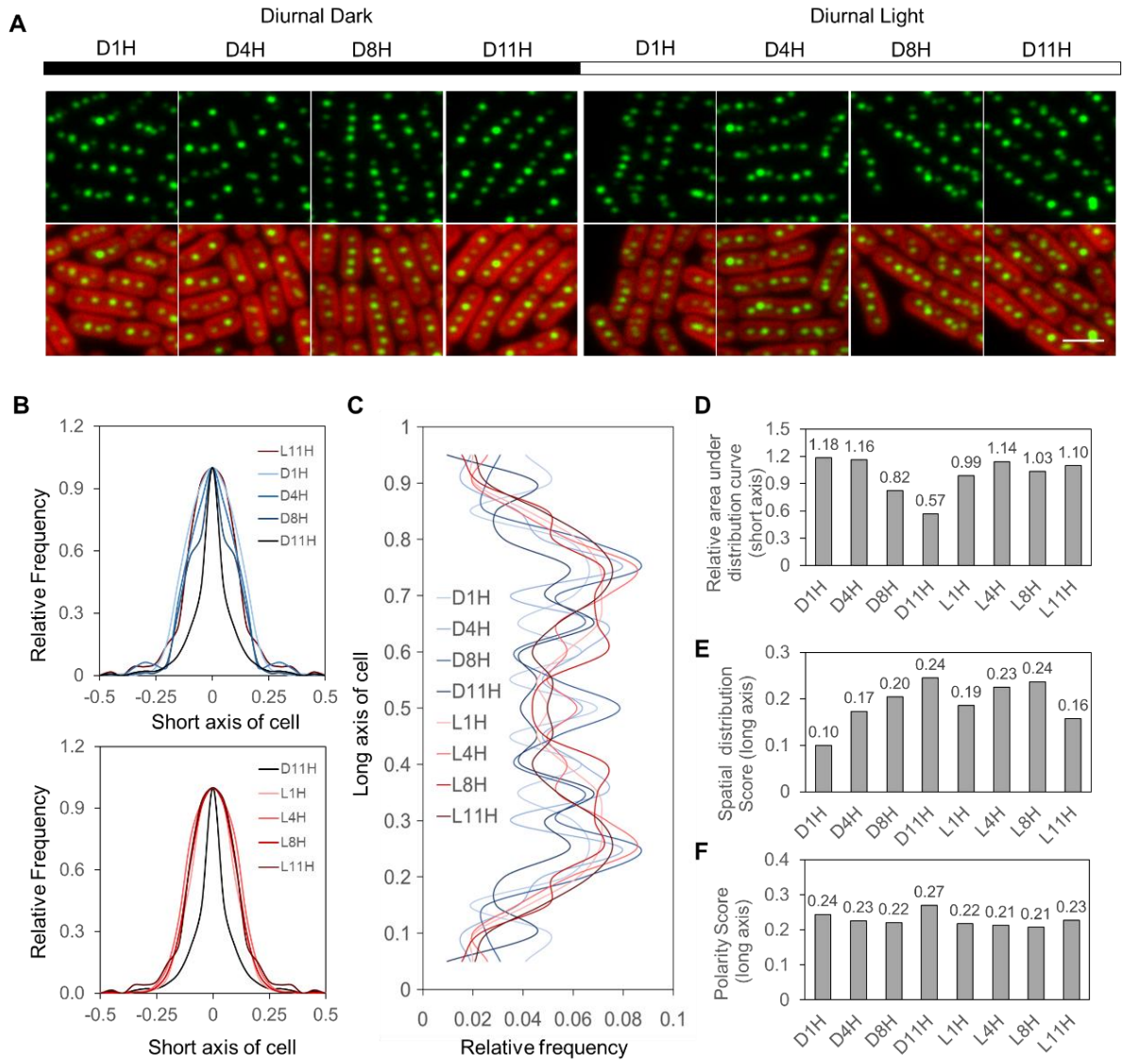


Figure 5-5. Carboxysome localisation in RbcL-eYFP cells under diurnal dark light conditions. A. representative confocal images taken at respective time points. **B.** Relative distribution frequencies on the short axis of the cell, a centralised to mid-axis could be observed during diurnal dark at D11H. **C.** Relative distribution frequencies on the long axis of the cell. **D.** Relative area under the curve for B for quantitative comparison and visualisation of positioning differences. **E.** Lower random distribution score over long axis indicating more random-like distribution (less spatial control) during dawn and initial hours in the dark (L11H to D1H). **F.** Polarity Score over long axis indicating no significant polar or midpoint positioning preference during diurnal light and dark condition. Equations for calculating random distribution score and polarity score are described in **Chapter 2** under **section 2.6.3**. For each time point, a minimum of 300 cells was analysed. The system errors for **D**, **E**, **and F** are 0.013, 0.011 and 0.014 respectively. Differences lesser than system errors were not considered as significant.

5.2.2 Rhythmic profile of carbon fixation capacity under diurnal condition

As demonstrated in Ch**apter 3**, CO_2 fixation capacity of carboxysome is strongly correlated with the abundance of Rubisco in cells. To evaluate CO_2 fixation capacity of carboxysomes during the diurnal conditions, we monitor the Rubisco abundance (**Figure 5-5A**, represented by fluorescent-tagged RbcL) and carbon fixation capacity through quantitative fluorescence microscopy and ^{14}C fixation assay (**Figure 5-6A)**. Repetitive tendencies in carbon fixation capacities were observed: the CO_2 fixation capacity continuously dropped during the dark period from D1H to D11H. The CO_2 fixation capacity was rescued after entering the light period (L1H) and reaches the highest in the day period at L4H and then decreases again from the 2nd half of light period at L8H and L11H. Overall, the averaging carbon fixation capacities during light period (5.0 ± 0.5 nmol·ml^{-1}·min^{-1}, n=15) were significantly higher (two-sided t-test P = 0.02) than those of the dark period (4.1 ± 0.6 nmol·ml^{-1}·min^{-1}, n=15). The carboxysome number per cell and RbcL-eYFP fluorescence intensities (representing the Rubisco content in the carboxysome) were also extracted from confocal images taken during the diurnal cycles. The carboxysome number per cell during L1H, L4H and L8H are significantly higher during at 4.1 ± 1.9 (n=4), 4.1 ± 2.2 (n=4) and 3.9 ± 2.0 (n=4) carboxysome per cell respectively comparing to the average of remaining time points (D1H, D4H, D8H, D11H, and L11H) at 3.4 ± 1.4 (n=4) carboxysome per cell (P<0.05), shown in **Figure 5-4B**. Interestingly, The Rubisco contents on each carboxysome were found in opposite fashion with the carboxysome number per cell: the lowest contents were observed in L1H, L4H and L8H at 0.81 ± 0.55 (n>500), 0.84 ± 0.58 (n>500), 0.85 ± 0.53 (n>500) (**Figure 5-4C**), compared with the rest time points (0.93 ± 0.61 (n>500), 1.00 ± 0.74 (n>500), 0.94 ± 0.59 (n>500), 0.98 ± 0.68 (n>500), and 0.94 ± 0.59 (n>500) as peak value ± HWHM (following method applied in **Chapter 4**) for D1H, D4H, D8H, D11H and L11H respectively). On the other hand, no significant differences (**Figure 5-4D**) of the overall Rubisco content, represented by the RbcL-eYFP signal, in the whole cell were detected (P >0.05, n=12 for the dark and light period).

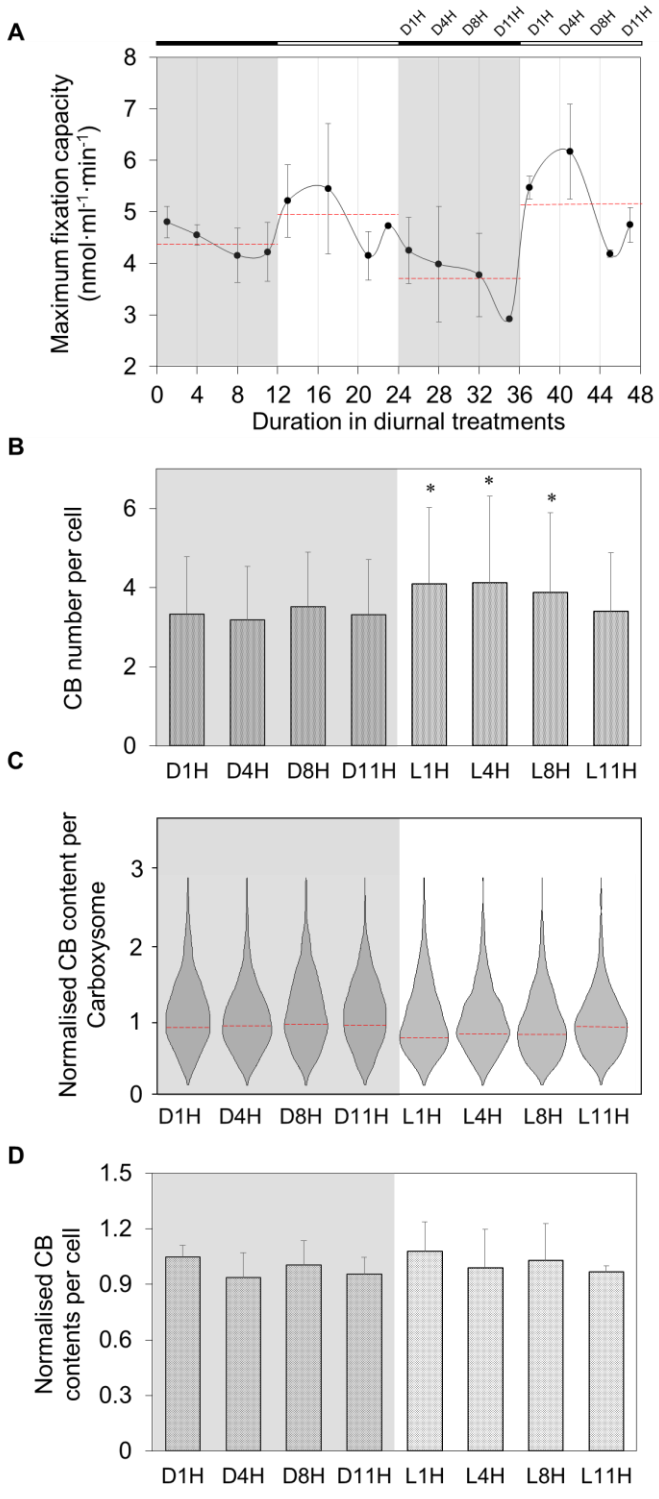


Figure 5-6. CO_2 fixation capacities, the carboxysome number and Rubisco loading per carboxysome and total Rubisco content within cell under diurnal dark light cycle. **A.** ^{14}C fixation assay for cells grown under diurnal 12D12L cycles. OD_{750} cell density was used for normalisation. The black-white bar above and grey-white background indicate the dark and light cycles respectively. The dotted lines with red colour indicate the average of fixation rate for cells in dark and light respectively. Error bar indicate SD from 3 biological repeats. **B-D.** Corresponding carboxysome number per cell, Rubisco content per carboxysome and total Rubisco content per cell (estimated by RbcL-eYFP content from fluorescence microscopy) during 2nd dark-light cycle in A. Error bars represent the SD from a minimum pool of 300 carboxysomes in **B**, and SD within three 3 different batch of biological repeats with a minimum of 40 cells in each repeat in **D**. Violin plots were generated by R illustrate the fluorescence intensity distribution of RbcL-eYFP during selected time points. The representative values and deviations were represented by Peak value from kernel density fitting and half width at half maximum (HWHM) Significances with *p*-value < 0.05 were marked by *.

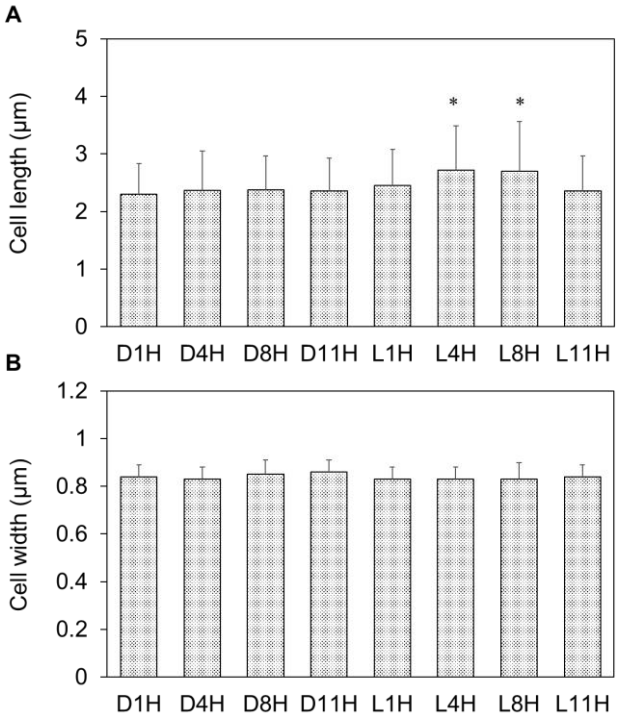


Figure 5-7. Cell dimensions of Syn7942 cells during diurnal condition. A. Averaging cell lengths are similar except during L4H and L8H ($P > 0.05$, $n > 300$ for each time points), while the cell widths during L4H and L8H are still within 10 % of the difference from the average widths. **B.** Difference in the averaged cell widths is not detectable ($P > 0.05$, $n = 300$). Error bars represent SD. Significance was marked as *.

Cell length and width were recorded during selected time points (**Figure 5-7**). No significant differences in the cell width were detected over the entire time course. We found less than 10 % increase in cell length at L4H (2.7 ± 0.8, n>300) and L8H (2.7 ± 0.9, n>300) compared to the rest time points (from D1H to D11H, L1H and L11H at 2.3 ± 0.5, 2.4 ± 0.7, 2.4 ± 0.6, 2.4 ± 0.6, 2.5 ± 0.6 and 2.4 ± 0.6 respectively, Mean ± SD, $n > 300$ for each time points). The increase in the cell length during L4H and L8H might be a combinatory outcome of cell elongation during the light period and preferable division close to dawn (Yang et al., 2010). Still, less than 10% increase in cell length was correlated with the 16-23% increase of carboxysome number per cell during corresponding L4H and L8H but unmatched to L1H in **Figure 5-6B**. At L1H, the increase of carboxysome number per cell from 3.3 ± 1.4 to 4.1 ± 1.9 appeared before elongation of cells observed at L4H.

While the transcriptomic data that categorise carboxysomal genes as class II gene with expression reach peak levels at dawn (Vijayan et al., 2009, Ito et al., 2009) corresponding to the D10H to D12H in our sampling time frame, no detectable increase in carbon fixation capacity, the carboxysome number as well as total Rubisco content per cell at D11H were found in our data, shown in **Figure 5-6A to C**. Instead, the highest Rubisco fixation capacity as well as the carboxysome number per cell were observed in light period during L1H, after 1-3 hours of the peak expression at the end of dark period. The unmatched profile further suggested the involvement of post-transcriptional regulation in carboxysome biogenesis, inconsistent with previous work (Guerreiro et al., 2014).

5.2.3 Diurnal profile for carbon fixation capacity could not be sustained under constant light

Circadian rhythms-controlled processes should keep persistence of rhythmic profile without external cues. We examined whether the rhythmic profile in carbon fixation capacity observed under diurnal conditions would remain under constant light (CL) conditions (**Figure 5-8A**). Remarkably, the iconic increase of carbon fixation capacity during L1H and L4H were not observed in both subjective light period at 13 h/15 h and 37 h/40 h corresponding to diurnal L1H/L4H. Instead, similar levels of CO_2 fixation compared with the previous subjective dark period from 0 h -12 h and 24 h -36 h were recorded. Meanwhile, we could still observe a similar reduction for CO_2 fixation capacity during diurnal light-dark transition across -36 h and -12 h during the subjective dark period from 0 h -12 h and 24 h -36 h, corresponding to diurnal D11H to L1H. To summarise, we observed similar decreasing tendencies of Rubisco fixation capacities in the subjective dark period between 0 h -12 h and 24 h -36 h while the increasing and peaking of fixation when entering the light period were dampened in the subjective light period during 12 h - 24 h and 36-48 h under CL condition. The persistent down-regulation of CO_2 fixation capacities during subjective dark period indicate possible negative regulation from circadian clock regardless the absence of darkness, while the positive up-regulation might mainly answering to the stimuli from dark-light shifting, therefore missing such stimulation under constant light lead a dampened up-regulation during subjective light period.

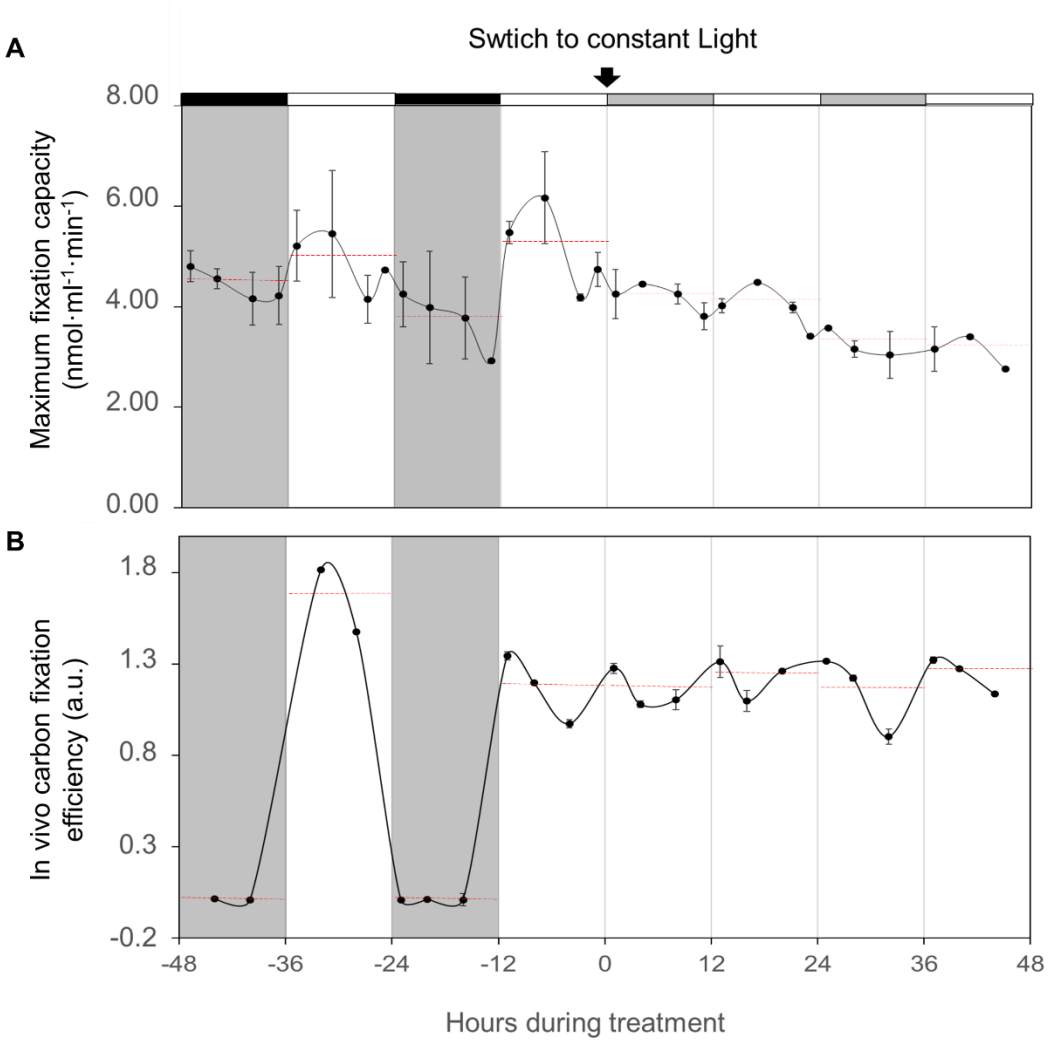


Figure 5-8. Comparison of carbon fixation capacities and in vivo carbon fixation efficiencies under diurnal/CL conditions in WT Syn7942. A. ^{14}C fixation capacity (shown as mean ± SD, n=3) for permeabilised cells grown under two continuous diurnal 12D12L cycles (-48h to 0h) and additional two days (0h to 48h) in CL. **B.** *In vivo* ^{14}C fixation efficiency (shown as mean ± SD, n=2) detected during growth in BG-11 medium with supplied low dosage of $NaH^{14}CO_3$ under diurnal and CL. Cell contents are normalised by cell density inferred through OD_{750} readings. Relative fixation rates were displayed in arbitrary unit (a.u.). The black-white-grey bar above indicate the dark, light, and subjective dark in light respectively. The red dotted lines indicate the average of sampled time points in corresponding 12-hour period respectively.

We also performed CO_2 fixation assays of living cells grown in BG-11 under both the diurnal condition and the subjective CL condition (described in **section 2.4.3**) to detect the real-time carbon fixation rates at corresponding time points that were measured for fixation capacities. Unlike the CO_2

fixation capacities showed in **Figure 5-8A** which was done with externally provided RuBP and bicarbonate at saturated concentration, *in vivo* CO_2 fixation assay was performed with endogenous RuBP and bicarbonate. Therefore, differences in the RuBP and bicarbonate contents between selected time points would also affect overall CO_2 fixation rate besides the impacts solely from the CO_2 fixation capacity of carboxysomes. We detected a 100-fold decrease in CO_2 fixation rate during both diurnal dark periods from -48 h to -36 h and -24 h to -12 h **(Figure 5-8B)**. The dramatic decreases were not detected during the subjective dark period from 0 h to 12 h and 24 h to 36 h. Instead, the same levels of fixation rate were observed. Combining the above data, we suggest that less than one-fold decrease in the carbon fixation capacity of carboxysome in diurnal dark period should not account for the drastic reduction of CO_2 fixation activity. Instead, the availability of substrate RuBP and bicarbonate might be the main rate-limiting factors.

5.2.4 Investigating circadian regulation in carboxysome biogenesis through ΔKaiA

We further investigate the **circadian regulation** by the circadian nulled Syn7942 strains (ΔKaiA and ΔKaiA::RbcL-eYFP) under constant moderate light (ML) at 50 $\mu E \cdot m^{-2}\ s^{-1}$ and circadian activated strains (WT and RbcL-eYFP) cells through knock out strategy. Due to the persistence of Syn7942 circadian clock (Cohen and Golden, 2015), we reset the clock of seeding cultures for the reference strains through two complete diurnal treatments, adapted in CL for an additional four days, and then used for inoculation in constant ML treatments. Confocal images taken for circadian null mutants and reference cells with an internal clock corresponding to the D1H time point were shown in **Figure 5-9A**. At a glance, the carboxysomes were visualised as typical multiple spot-like signals in both ΔKaiA::RbcL-eYFP and RbcL-eYFP strains, indicating no major disruption of carboxysome formation caused by KaiA nullification. We further compared the distribution of carboxysome in ΔKaiA::RbcL-eYFP against RbcL-eYFP illustrated as heat maps (**Figure 5-9B**) and distribution profiles along the short and long axis of the cells (**Figure 5-9C**). The carboxysome distribution along the short axis of cell quantified as the area under the distribution curve indicate no pronounced cell-central preference in both ΔKaiA and control Syn7942 strains (**Figure 5-9C)** (Area under peak = 1.18

and 0.99 for RbcL-eYFP and ΔKaiA::RbcL-eYFP respectively). The value for the area under peak for RbcL-eYFP matched well with the corresponding value at D1H in diurnal treatments (Also at 1.18, **Figure 5-5D**). Meanwhile, we observed noticeable differences in the distribution profiles along the long axis of the cell. In RbcL-eYFP, two hot zones at the cell poles and two additional hot zones at quarter locations were found (illustrated as the grey shading area in **Figure 5-9C**) similar to the case for RbcL-eGFP in Chapter 3, while the two hot zones at cell quarter were less pronounced in the ΔKaiA::RbcL-eYFP strain. The spatial distribution scores for RbcL-eYFP is 0.08, closely resemble the score under diurnal condition at corresponding D1H at 0.10 (**Figure 5-5E**). While in ΔKaiA::RbcL-eYFP a higher score at 0.13 might suggest a stronger positioning control during D1H, which was loosened with the presence of circadian control (**Figure 5-5E**). The Polarity scores are 0.24 and 0.23 for RbcL-eYFP and ΔKaiA::RbcL-eYFP respectively, both identical to the value at 0.24 calculated for RbcL-eYFP during D1H under diurnal condition (**Figure 5-5F**). Meanwhile, higher carboxysome number per cell was observed at 4.64 ± 1.09 in ΔKaiA::RbcL-eYFP mutant compared to the RbcL-eYFP control at 3.07 ± 0.76. (**Figure 5-9D)**. A strikingly near 1-fold decrease in YFP fluorescence content per carboxysome was observed in ΔKaiA::RbcL-eYFP (**Figure 5-9E**). The reduced Rubisco content on each carboxysome and increased carboxysome number per cell in ΔKaiA background suggested important roles played by circadian in cellular regulation of carboxysome biogenesis.

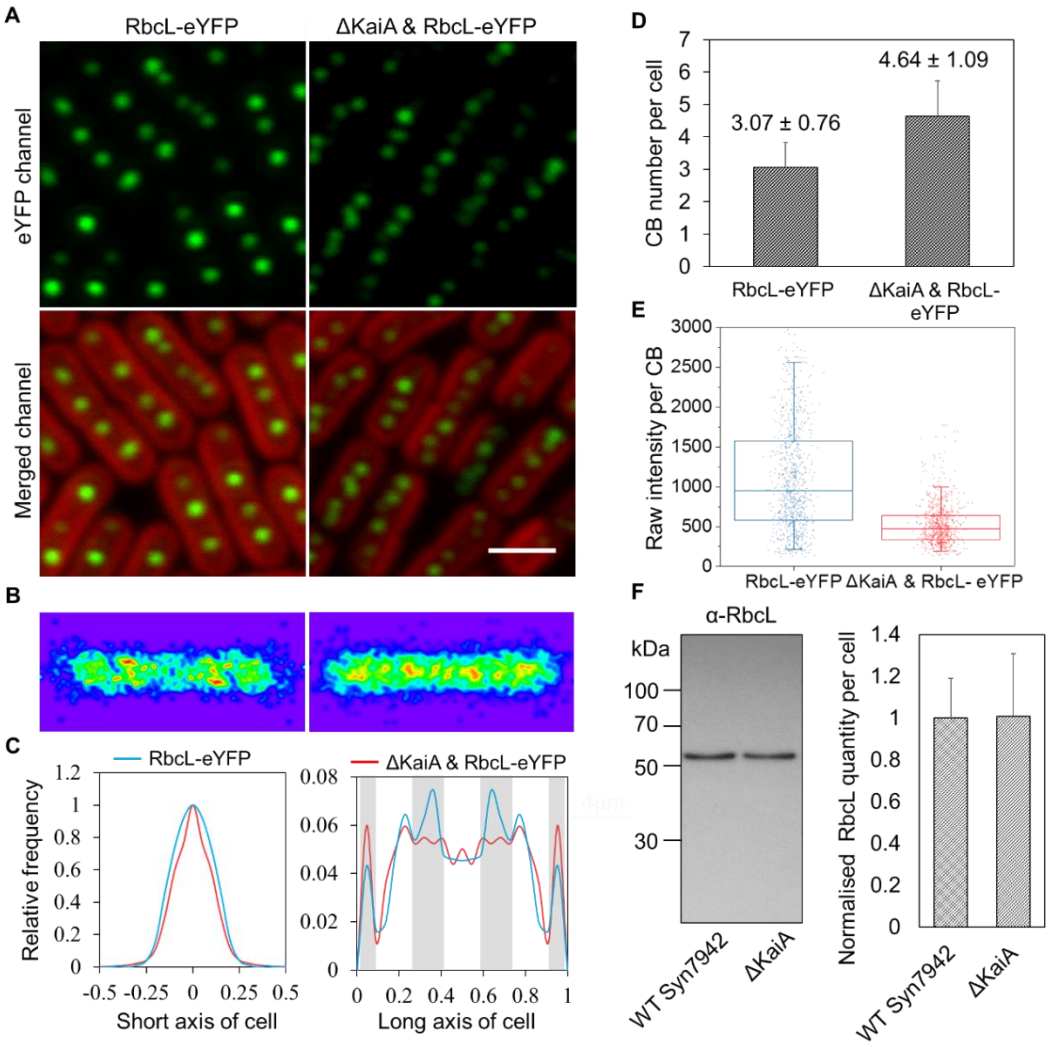


Figure 5-9. Biogenesis of carboxysome in cells is alternated in ΔKaiA background under constant light condition. A. Confocal images for RbcL-eYFP and ΔKaiA::RbcL-eYFP under constant light (50 μE·m^{-2} s^{-1}), images are enhanced for visualisation purposes; all signals capture in the analysis were below saturation. Scale bar=2 μm. **B** and **C.** Heat map and carboxysome distribution profiles (red for ΔKaiA) indicating unaltered carboxysome localisation along the short cell axis and altered distribution over the long axis of the cell (cell number>300 for each strain). The hot spots for carboxysome localisation in RbcL-eYFP were illustrated as grey background. **D.** carboxysome number per cell measured from confocal images indicate increased carboxysome number in ΔKaiA compared with control (carboxysome number >300 for each strain). **E.** Signal quantification for each carboxysome of RbcL-eYFP in WT and KaiA null background indicating reduced Rubisco loads per carboxysome result from ΔKaiA (carboxysome number >300 for each strain). **F.** Quantification of RbcL by immunoblotting with an anti-RbcL antibody in WT and ΔKaiA indicates same levels of Rubisco contents at the cellular level caused by KaiA depletion. Representative Immunoblot visualisation images and quantification from 3 biological repeats were shown as average ± SD.

To confirm the cellular level of Rubisco in the cells, we quantify the Rubisco content per cell by immunoblotting with anti-RbcL antibody in Syn7942 Δ*kaiA* and WT strains that grown under same constant light condition, sampled at the same time with the YFP mutants (**Figure 5-9F)**. The quantification indicates equal levels of overall Rubisco contents per cell with or without KaiA under constant light (**Figure 5-9F)**. The fixation capacities were also unaltered inferred from carbon fixation assay data (**Figure 5-10**).

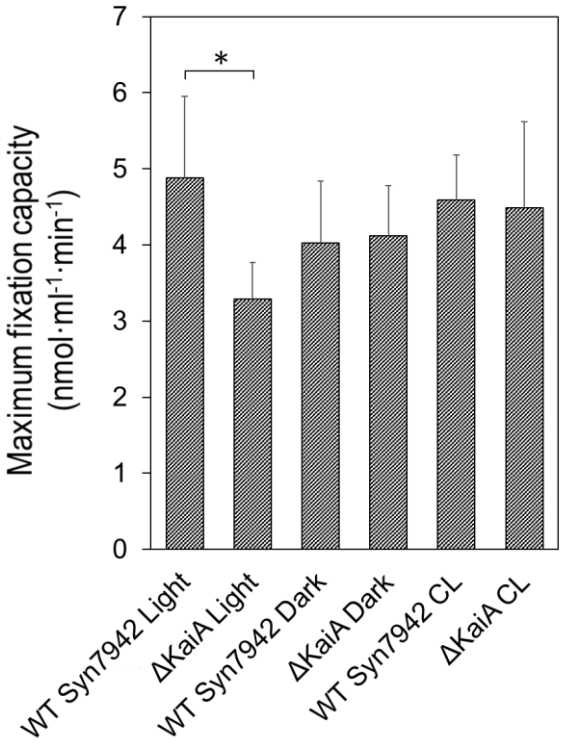


Figure 5-10. ^{14}C fixation assay for WT Syn7942 and ΔKaiA during diurnal light, dark and constant light (CL) conditions. Reduced fixation capacity was found in ΔKaiA cells compared with WT during diurnal light period ($p<0.05$). Error bar represents SD from 3 biological samples.

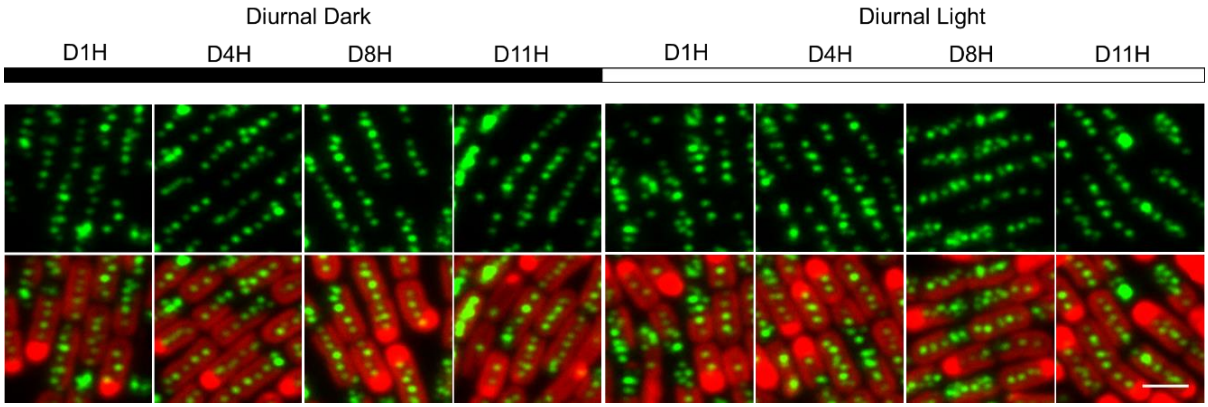


Figure 5-11. The carboxysome localisation in ΔKaiA::RbcL-eYFP cells under diurnal dark light conditions. Typical confocal images taken at respective time points. Due to the irregularity of chlorophyll signals, cell recognition cannot be done by Image SXM to evaluate carboxysome localisation within cells.

We then applied the similar diurnal treatment on the ΔKaiA::RbcL-eYFP mutant to address the diurnal light effect without the functional circadian core oscillator, shown in **Figure 5-11**. However, unlike CL condition, significant disruption to chlorophyll was observed under the diurnal condition, suggesting irregular photosynthesis caused by missing of KaiA. As a result, we could not perform further quantification by pipelined software through Image SXM due to the false cell recognition. We then performed carbon fixation assay to determine the functionality of carboxysomes in the ΔKaiA compared with WT that were both grown under constant light and diurnal condition. The averaging fixation capacities shown in **Figure 5-10** indicate significantly decreased of fixation capacity in ΔKaiA during the light period of diurnal condition (p<0.05), suggesting severe impact caused by KaiA knockout in carbon fixation capacities. Meanwhile, no significant difference was detected during the dark period of the diurnal condition (**Figure 5-10)**.

5.3 Discussion

In this chapter, we monitored the carboxysome localisation in Syn7942 during diurnal condition. The delayed linear-like centralised repositioning of carboxysomes on the short axis of cell stimulated by the electron transport inhibitor DBMIB was initially reported in **Chapter 3**, believed to be correlated with a reduced redox state of PQ pool (Sun et al., 2016). The redox state of photosynthetic electron

transport chain play essential roles in the distribution of respiratory complexes (Liu et al., 2012), photosystem composition (Fujita et al., 1987), photosynthetic state transitions (Mullineaux and Allen, 1990) and the modulation of the circadian clock (Ivleva et al., 2006; Wood et al., 2010). In nature, the PQ pool is periodically shifted from an oxidised state to reduced state during light and dark respectively (Schuurmans et al., 2014). The fact that such centralising rearrangement, can be stimulated not only by externally provided PQ pool manipulator, but also through exposure in natural day-night condition further confirms the natural linkage between redox state of PQ pool with carboxysome repositioning. The delayed pattern of localisation changes does not resemble typical circadian controlled phenotypes as no changes were made before shifting of light and dark. We assume the carboxysome distribution over the short axis of the cell should be a result of answering to the redox state instead of regulation by circadian.

Meanwhile, correlation of circadian control in the arrangement of carboxysomes along the long axis of the cell could be established, inferred from the reposition before light-dark shifting. We assumed the decrease of spatial distribution score before entering dark period at L11H might not be an outcome merely stimulate through the sensing of light. Instead, such "prepare-ahead" strategy might suggest the involvement of circadian regulation. The reducing control of carboxysome localisation along the cell long axis during dusk was coped with a lower CO_2 fixation capacity and *in vivo* fixation rate (**Figure 5-8A** and **B**), together with other circadian controlled activities such as cell division (Mori et al., 1996; Yang et al., 2010) and chromosome compactions (Smith and Williams, 2006). The positioning of chromosomes and carboxysomes was reported to be mutual exclusively localised even when carboxysome localisation was disrupted through the deletion of *parA* (Jain et al., 2012). The suggested independencies of positioning mechanisms between chromosomes and carboxysomes might indicate some direct and active regulation in carboxysomes partitioning along the cell rather than passively altered through reorganisation of chromosomes. The recently discovered vital component, McdA, and McdB that crucial for carboxysome localisation (MacCready et al., 2018) might be worthy to be examined under diurnal conditions in future to investigate further the purposes and functions of carboxysome repositioning in cells.

Moreover, we could fuse fluorescent protein to Kai protein to monitor the aggregation and disassociation of complexes that was reported (Cohen et al., 2014) while visualising the carboxysome, achieved in **Figure 5-12A**. However, we observed multiple KaiA aggregates during D8H in one cell instead of one that was located at the cell pole. Nevertheless, to better illustrate the dynamic changes in carboxysome positioning, more time points instead of 1, 4, 8, 11H in a 12-hour cycle should be taken. Therefore, time-lapse imaging might be superior to provide more details in such dynamic changes across the diurnal cycles. The feasibility of the time-lapse experiment was demonstrated in **Figure 5-12B** with a 10-minute interval.

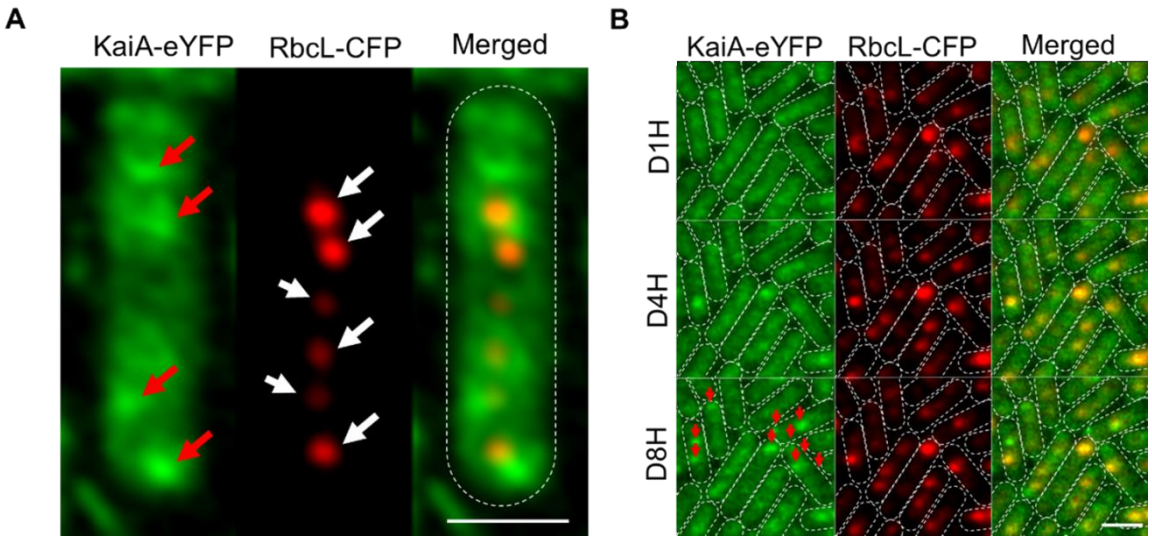


Figure 5-12. Fluorescence microscopy images of KaiA-eYFP::RbcL-CFP double mutant enable visualisation of carboxysomes and KaiA complex aggregation in the same view. A. Typical confocal images were taken at D8H shown both carboxysomes (marked by white arrows) and aggregated Kai complexes (marked by red arrows). Scale bar = 1 μm. **B.** time-lapse images are showing the formation of Kai complex during the dark period, marked by red arrows. Scale bar = 2 μm.

On the other hand, the rhythmic profile of carbon fixation capacity (**Figure 5-8A**) and relatively constant Rubisco content (**Figure 5-6D)** might suggest alternated CO_2 fixation efficiency per Rubisco throughout the diurnal cycles, though estimation for such unaltered Rubisco abundancies in cells though a more precise method, ideally quantification through immunoblotting by anti-RbcL antibody is necessary. The CO_2 fixation efficiency per Rubisco can be regulated through the circadian oscillated

expression of Rubisco activase in higher plants (Liu et al., 1996; To et al., 1999), or through the repositioning of Rubisco in a chloroplast-containing marine algae Dinoflagellate (Nassoury et al., 2001). In the CO_2 fixation capacity measurement (**Figure 5-6A** and **Figure 5-8A**), the cell samples went through a permeabilisation process that puncture holes on cell membranes, allowing provided RuBP and bicarbonate in assay buffer to enter the cell cytosol. Therefore, the differences in carboxysome localisation within the cells would not be factored into the fixation capacity measured. Instead, structural and functional modification through dynamical relocation of shell proteins over formed proteins sheets (Sutter et al., 2016) that allows permeability tuning or stoichiometry tuning that answer to different levels of external stimuli such as light and CO_2 (discussed in chapter 4) might play crucial roles. Quantification and localisation for other carboxysomal proteins such as CcmK2-K4, CcmL and CcaA combining RbcL that indicate the stoichiometry alternation on carboxysome might provide further evidence for the purposed assumption.

The pore of shell proteins was capable of changing permeability (open and close) through alternative side-chain conformations, reported in BMC-T protein CcmP (Cai et al., 2013). Even though it remains unclear whether the conformation changes could happen *in vivo*, and which particular stimuli that trigger such conformation changes, we could not rule out the possibility for the differentially open/close gateway mechanism coping with the dark and light conditions. Particularly when CcmP nullification mutant process near WT phenotype when evaluated under constant light conditions (data not included in this thesis). Combining CcmP knockout strains in future works might provide a more definite answer to whether gateway system exists under diurnal conditions.

Initially, ΔKaiA was constructed with a purpose to evaluate solely the circadian null phenotype, which was commonly adopted (Clerico et al., 2009). A significant decreased in CO_2 fixation efficiency per Rubisco was found in ΔKaiA during the light period of the cyclic light (**Figure 5-10**). However, recent work has shown severe damage to diurnal growth caused by KaiA knockout compared to entirely KaiABC knockout. Suggesting additional roles was suggested for KaiA (Welkie et al., 2018). Therefore, the difference we observed in ΔKaiA::RbcL-eYFP might not solely come

from the nullification of the circadian clock. Future work with KaiABC knockout was required to confirm the exact source of observed phenotype.

In **Chapter 3** and **4**, we investigate the modulation of carboxysome in correspondence with the light intensities. In this chapter, we documented the carboxysome biogenesis under cyclic light, further revealing the comprehensive regulation that was embedded into the entire metabolic systems in host cells. In the author's view, a better understanding of how carboxysome biogenesis in integrated into the native metabolic pathway is crucial for carboxysome-based bioengineering. For instance, introducing of carboxysomes into plants, where circadian and diurnal regulation is strong and dominant.

Chapter 6

Conclusions and Perspectives

Chapter 6

6.1 Conclusions

This thesis has provided new information to improve our understanding on the fine-tuning of carboxysome biogenesis (content quantity, stoichiometry and assembly, as well as subcellular localisations) under native regulatory system in responding to changing environmental conditions such as light intensities (**Chapter 3** and **4**), light patterns (**Chapter 5**) as well as CO_2 availability (**Chapter 5**).

Before this work, little was known about how carboxysome biosynthesis and spatial positioning are physiologically regulated to adjust to dynamic changes in the environment. We demonstrated that β-carboxysome biosynthesis is enhanced through increased unit counts per cell responding to increasing light intensity, thereby enhancing the carbon fixation activity of the cell. Inhibition of photosynthetic electron flow impairs the accumulation of carboxysomes, indicating close coordination between β-carboxysome biogenesis and photosynthetic electron transport. Likewise, the spatial organisation of carboxysomes in the cell correlates with the redox state of photosynthetic electron transport chain. This study provides essential knowledge for us to modulate the β-carboxysome biosynthesis and function in cyanobacteria. In translational terms, the knowledge is instrumental for design and synthetic engineering of functional carboxysomes into higher plants to improve photosynthesis performance and CO_2 fixation.

Carboxysomes are complex, proteinaceous organelles that play essential roles in carbon assimilation in cyanobacteria and chemoautotrophs, comprising hundreds of self-assembled protein homologs. Despite their significance in enhancing CO_2 fixation and great potential in bioengineering applications, their structural composition and adaptation to cope with environmental stress remain unclear. In this work, we find that the protein stoichiometry, diameter, cellular localisation and mobility pattern of carboxysomes in cells depend sensitively on the microenvironmental levels of CO_2 and light intensity during cell growth, revealing cellular strategies of dynamic regulation. To our knowledge, this is the first evidence describing the absolute numbers of proteins accurately on individual carboxysome, at

the single-organelle level. A deeper understanding of the molecular composition and regulation of carboxysomes is fundamental to rational design and construction of functional carboxysomes in heterologous organisms using synthetic biology, with the intent of enhancing cellular metabolism and biomass production.

In nature, light accessible to cyanobacteria in a diurnal fashion, unlike the constant light commonly used in laboratories. Before this work, it is unclear whether and how diurnal light and the circadian clock mediate carboxysome biogenesis in natural habitats. In this work, we inspect carboxysome biogenesis under near-nature conditions with a day-night cycle and strengthen the possible correlation of carboxysome localisation regulation naturally occurred with the redox state of the plastoquinone pool (PQ pool) under diurnal condition. The rhythmic carbon fixation capacities during the light period of diurnal cycles were quickly dampened after switching to constant light suggest dominant influence of light illuminations on carbon fixation activities. Lastly, we inspected the carboxysome numbers, localisation and enzymatic properties in a circadian nullified mutant ΔKaiA. The results suggested changed carboxysome numbers per cell and Rubisco content per carboxysome, whereas relatively constant cellular level fixation capacity and Rubisco content. This study provides un-documented localisation, enzymatic activity profile of carboxysomes in the circadian-nullified background for future works undertaking circadian regulation in carboxysome biogenesis. The findings suggesting rhythmic tunings of carboxysome biogenesis coupling with diurnal day-night cycles in native cells and emphasise the importance of considering diurnal regulation in future works aiming at integrating and utilising carboxysomes heterologously.

6.2 Perspectives

In **Chapter 3**, we reported the modulation of carboxysomes in response to light intensities changes and elucidated the electron flow as driven force to modified biogenesis of carboxysomes. Recently, a step has been taken forward to further dissecting the light sensing to specific signalling pathway through photoreceptor RcaE (Rohnke et al., 2018). The RcaE controls the tuning of photosynthetic

efficiency during long-termed acclimation. The RcaE mediated light modulation therefore affects carboxysome biogenesis through a progressive process in long terms, which matched our findings suggesting adaptions of carboxysome contents that adapted in a duration of several days (**Figure 3-8**). Moreover, light quality as spectrum (red and green light) are further divided and changed stoichiometry of structural protein CcmL, CcmM, CcmN and CcmO are reported under different light spectrums. The findings further confirmed the tuning of structure and functions of carboxysomes in response to dynamic light illuminations. Future investigation for the triggering mechanism would be crucial to interlink light harvesting/sensing to carboxysome modulations.

Moreover, tuning of carboxysome structure and functions under environmental stresses other than light and CO_2 also worth future attention. Beyond the scope of this thesis, evidence of transcriptomic levels has suggest expressional variation of carboxysomal genes under nitrogen (Osanai et al., 2006) and oxidative (Kobayashi et al., 2004) stresses. Nitrogen metabolism is important for protein synthesis and a carbon/nitrogen balancing is purposed to have carboxysome biogenesis in pace with nitrogen assimilations. Future works investigating cells under these stresses might reveal missing regulation that link multiple biological processes cross-talking to carboxysome biogenesis.

Photosynthetic activity is a strong input to the cellular redox and carbon fixations by carboxysomes might be in the core position rebalancing the redox homeostasis. In **Chapter 3** and **5**, we've found the localisation of carboxysomes in cells in strong correlation with redox states (**Figure 3-14**, **Figure 5-5**). Future works investigating carboxysome biogenesis under oxidative stress such as exposure to methyl viologen besides HL might yield more informative findings regarding the possible hidden roles of redox balancing by carboxysomes.

With the strains generated and time-lapse imaging for long-term carboxysomal signal tracking as well as carboxysome subcellular localisation analysis toolbox developed in this project (described in

Chapter 2), we could further investigate carboxysome modulations in dynamical fashions to aid in the interpretation of circadian and diurnal light regulation documented in **Chapter 5**. Calvin cycle is regulated by CP12 protein via the formation of PRK/GADPH/CP12 complex (Tamoi et al., 2005), future work looking into the role of CP12 in context of carboxysome biogenesis would be crucial to understand the details of hidden regulatory under diurnal light condition. Moreover, the available tools enable us to investigate carboxysome behaviour during cell divisions. To date, very limited information is available regarding the mobilisation of carboxysomes particularly during cell divisions (Chen et al., 2013). Combining dual labelling fluorescent microscopy together with a series of knockout strains generated over individual carboxysomal proteins (data not included in this thesis), we could further investigate stoichiometry and localisation in real time to determine the pace alternation during alternated environmental conditions.

Reference

Agarwal, R., Ortleb, S., Sainis, J.K., and Melzer, M. (2009). Immunoelectron microscopy for locating calvin cycle enzymes in the thylakoids of synechocystis 6803. Mol Plant *2*, 32-42.

Agirrezabala, X., Méndez-López, E., Lasso, G., Sánchez-Pina, M.A., Aranda, M., and Valle, M. (2015). The near-atomic cryoEM structure of a flexible filamentous plant virus shows homology of its coat protein with nucleoproteins of animal viruses. eLife *4*, e11795.

Allahverdiyeva, Y., Suorsa, M., Tikkanen, M., and Aro, E.M. (2015). Photoprotection of photosystems in fluctuating light intensities. J Exp Bot *66*, 2427-2436.

Aryal, U.K., Stöckel, J., Krovvidi, R.K., Gritsenko, M.A., Monroe, M.E., Moore, R.J., Koppenaal, D.W., Smith, R.D., Pakrasi, H.B., and Jacobs, J.M. (2011). Dynamic proteomic profiling of a unicellular cyanobacterium Cyanothece ATCC51142 across light-dark diurnal cycles. BMC Systems Biology *5*, 194-194.

Asayama, M. (2006). Regulatory system for light-responsive gene expression in photosynthesizing bacteria: cis-elements and trans-acting factors in transcription and post-transcription. Biosci Biotechnol Biochem *70*, 565-573.

Aussignargues, C., Paasch, B.C., Gonzalez-Esquer, R., Erbilgin, O., and Kerfeld, C.A. (2015). Bacterial microcompartment assembly: The key role of encapsulation peptides. Communicative & Integrative Biology *8*, e1039755.

Axen, S.D., Erbilgin, O., and Kerfeld, C.A. (2014). A taxonomy of bacterial microcompartment loci constructed by a novel scoring method. PLoS Comput Biol *10*, e1003898.

Badger, M.R., and Price, G.D. (2003). CO2 concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution. Journal of Experimental Botany *54*, 609-622.

Badger, M.R., Price, G.D., Long, B.M., and Woodger, F.J. (2006). The environmental plasticity and ecological genomics of the cyanobacterial CO2 concentrating mechanism. J Exp Bot *57*, 249-265.

Badrinarayanan, A., Reyes-Lamothe, R., Uphoff, S., Leake, M.C., and Sherratt, D.J. (2012). In vivo architecture and action of bacterial structural maintenance of chromosome proteins. Science *338*, 528-531.

Bartolák-Suki, E., Imsirovic, J., Nishibori, Y., Krishnan, R., and Suki, B. (2017). Regulation of Mitochondrial Structure and Dynamics by the Cytoskeleton and Mechanical Factors. International Journal of Molecular Sciences *18*, 1812.

Baumgart, M., Huber, I., Abdollahzadeh, I., Gensch, T., and Frunzke, J. (2017). Heterologous expression of the Halothiobacillus neapolitanus carboxysomal gene cluster in Corynebacterium glutamicum. J Biotechnol *258*, 126-135.

Bobik, T.A. (2006). Polyhedral organelles compartmenting bacterial metabolic processes. Applied Microbiology and Biotechnology *70*, 517-525.

Bobik, T.A., Havemann, G.D., Busch, R.J., Williams, D.S., and Aldrich, H.C. (1999). The propanediol utilization (pdu) operon of Salmonella enterica serovar Typhimurium LT2 includes genes necessary for formation of polyhedral organelles involved in coenzyme B(12)-dependent 1, 2-propanediol degradation. J Bacteriol *181*, 5967-5975.

Bobik, T.A., Lehman, B.P., and Yeates, T.O. (2015). Bacterial microcompartments: widespread prokaryotic organelles for isolation and optimization of metabolic pathways. Molecular microbiology *98*, 193-207.

Bonacci, W., Teng, P.K., Afonso, B., Niederholtmeyer, H., Grob, P., Silver, P.A., and Savage, D.F. (2012). Modularity of a carbon-fixing protein organelle. Proc Natl Acad Sci U S A *109*, 478-483.

Boxall, S.F., Dever, L.V., Knerova, J., Gould, P.D., and Hartwell, J. (2017). Phosphorylation of Phosphoenolpyruvate Carboxylase Is Essential for Maximal and Sustained Dark CO2 Fixation and Core Circadian Clock Operation in the Obligate Crassulacean Acid Metabolism Species Kalanchoe fedtschenkoi. Plant Cell *29*, 2519-2536.

Bracher, A., Starling-Windhof, A., Hartl, F.U., and Hayer-Hartl, M. (2011). Crystal structure of a chaperone-bound assembly intermediate of form I Rubisco. Nature structural & molecular biology *18*, 875-880.

Bullerjahn, G.S., and Post, A.F. (2014). Physiology and molecular biology of aquatic cyanobacteria. Front Microbiol *5*, 359.

Burnap, R.L., Hagemann, M., and Kaplan, A. (2015). Regulation of CO(2) Concentrating Mechanism in Cyanobacteria. Life *5*, 348-371.

Burnap, R.L., Nambudiri, R., and Holland, S. (2013). Regulation of the carbon-concentrating mechanism in the cyanobacterium Synechocystis sp. PCC6803 in response to changing light intensity and inorganic carbon availability. Photosynthesis research *118*, 115-124.

Bustos, S.A., and Golden, S.S. (1992). Light-regulated expression of the psbD gene family in Synechococcus sp. strain PCC 7942: evidence for the role of duplicated psbD genes in cyanobacteria. Molecular & general genetics : MGG *232*, 221-230.

Cai, F., Bernstein, S.L., Wilson, S.C., and Kerfeld, C.A. (2016). Production and Characterization of Synthetic Carboxysome Shells with Incorporated Luminal Proteins. Plant Physiol *170*, 1868-1877.

Cai, F., Dou, Z., Bernstein, S.L., Leverenz, R., Williams, E.B., Heinhorst, S., Shively, J., Cannon, G.C., and Kerfeld, C.A. (2015a). Advances in Understanding Carboxysome Assembly in Prochlorococcus and Synechococcus Implicate CsoS2 as a Critical Component. Life (Basel) *5*, 1141-1171.

Cai, F., Kerfeld, C.A., and Sandh, G. (2012). Bioinformatic Identification and Structural Characterization of a New Carboxysome Shell Protein. In Functional Genomics and Evolution of Photosynthetic Systems, R. Burnap, and W. Vermaas, eds. (Dordrecht: Springer Netherlands), pp. 345-356.

Cai, F., Menon, B.B., Cannon, G.C., Curry, K.J., Shively, J.M., and Heinhorst, S. (2009). The pentameric vertex proteins are necessary for the icosahedral carboxysome shell to function as a CO2 leakage barrier. PLoS One *4*, e7521.

Cai, F., Sutter, M., Bernstein, S.L., Kinney, J.N., and Kerfeld, C.A. (2015b). Engineering bacterial microcompartment shells: chimeric shell proteins and chimeric carboxysome shells. ACS Synth Biol *4*, 444-453.

Cai, F., Sutter, M., Cameron, J.C., Stanley, D.N., Kinney, J.N., and Kerfeld, C.A. (2013). The structure of CcmP, a tandem bacterial microcompartment domain protein from the beta-carboxysome, forms a subcompartment within a microcompartment. J Biol Chem *288*, 16055-16063.

Cameron, J.C., Wilson, S.C., Bernstein, S.L., and Kerfeld, C.A. (2013). Biogenesis of a bacterial organelle: the carboxysome assembly pathway. Cell *155*, 1131-1140.

Chen, A.H., Robinson-Mosher, A., Savage, D.F., Silver, P.A., and Polka, J.K. (2013). The bacterial carbon-fixing organelle is formed by shell envelopment of preassembled cargo. PLoS One *8*, e76127.

Chen, P., Andersson, D.I., and Roth, J.R. (1994). The control region of the pdu/cob regulon in Salmonella typhimurium. J Bacteriol *176*, 5474-5482.

Chong, Y.T., Koh, J.L., Friesen, H., Duffy, S.K., Cox, M.J., Moses, A., Moffat, J., Boone, C., and Andrews, B.J. (2015). Yeast Proteome Dynamics from Single Cell Imaging and Automated Analysis. Cell *161*, 1413-1424.

Chowdhury, C., Chun, S., Pang, A., Sawaya, M.R., Sinha, S., Yeates, T.O., and Bobik, T.A. (2015). Selective molecular transport through the protein shell of a bacterial microcompartment organelle. Proceedings of the National Academy of Sciences *112*, 2990-2995.

Chowdhury, C., Sinha, S., Chun, S., Yeates, T.O., and Bobik, T.A. (2014). Diverse bacterial microcompartment organelles. Microbiol Mol Biol Rev *78*, 438-468.

Chuartzman, S.G., Nevo, R., Shimoni, E., Charuvi, D., Kiss, V., Ohad, I., Brumfeld, V., and Reich, Z. (2008). Thylakoid Membrane Remodeling during State Transitions in Arabidopsis. The Plant Cell *20*, 1029-1039.

Clerico, E.M., Cassone, V.M., and Golden, S.S. (2009). Stability and lability of circadian period of gene expression in the cyanobacterium Synechococcus elongatus. Microbiology *155*, 635-641.

Cohen, S.E., Erb, M.L., Pogliano, J., and Golden, S.S. (2015). Best practices for fluorescence microscopy of the cyanobacterial circadian clock. Methods Enzymol *551*, 211-221.

Cohen, S.E., Erb, M.L., Selimkhanov, J., Dong, G., Hasty, J., Pogliano, J., and Golden, S.S. (2014). Dynamic localization of the cyanobacterial circadian clock proteins. Current biology : CB *24*, 1836-1844.

Cohen, S.E., and Golden, S.S. (2015). Circadian Rhythms in Cyanobacteria. Microbiol Mol Biol Rev *79*, 373-385.

Crowley, C.S., Cascio, D., Sawaya, M.R., Kopstein, J.S., Bobik, T.A., and Yeates, T.O. (2010). Structural Insight into the Mechanisms of Transport across the Salmonella enterica Pdu Microcompartment Shell. The Journal of Biological Chemistry *285*, 37838-37846.

Crowley, C.S., Sawaya, M.R., Bobik, T.A., and Yeates, T.O. (2008). Structure of the PduU shell protein from the Pdu microcompartment of Salmonella. Structure *16*, 1324-1332.

Dai, W., Chen, M., Myers, C., Ludtke, S.J., Pettitt, B.M., King, J.A., Schmid, M.F., and Chiu, W. (2018). Visualizing Individual RuBisCO and Its Assembly into Carboxysomes in Marine Cyanobacteria by Cryo-Electron Tomography. Journal of Molecular Biology.

Dalla Chiesa, M., Friso, G., Deak, Z., Vass, I., Barber, J., and Nixon, P.J. (1997). Reduced turnover of the D1 polypeptide and photoactivation of electron transfer in novel herbicide resistant mutants of Synechocystis sp. PCC 6803. European journal of biochemistry *248*, 731-740.

Datsenko, K.A., and Wanner, B.L. (2000). One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc Natl Acad Sci U S A *97*, 6640-6645.

Dekker, J.P., and Boekema, E.J. (2005). Supramolecular organization of thylakoid membrane proteins in green plants. Biochimica et biophysica acta *1706*, 12-39.

Ditty, J.L., Canales, S.R., Anderson, B.E., Williams, S.B., and Golden, S.S. (2005). Stability of the Synechococcus elongatus PCC 7942 circadian clock under directed anti-phase expression of the kai genes. Microbiology *151*, 2605-2613.

Dodd, A.N., Salathia, N., Hall, A., Kevei, E., Toth, R., Nagy, F., Hibberd, J.M., Millar, A.J., and Webb, A.A. (2005). Plant circadian clocks increase photosynthesis, growth, survival, and competitive advantage. Science *309*, 630-633.

Dou, Z., Heinhorst, S., Williams, E.B., Murin, C.D., Shively, J.M., and Cannon, G.C. (2008). CO2 fixation kinetics of Halothiobacillus neapolitanus mutant carboxysomes lacking carbonic anhydrase suggest the shell acts as a diffusional barrier for CO2. J Biol Chem *283*, 10377-10384.

Drews, G., and Niklowitz, W. (1956). [Cytology of Cyanophycea. II. Centroplasm and granular inclusions of Phormidium uncinatum]. Archiv fur Mikrobiologie *24*, 147-162.

Dryden, K.A., Crowley, C.S., Tanaka, S., Yeates, T.O., and Yeager, M. (2009). Two-dimensional crystals of carboxysome shell proteins recapitulate the hexagonal packing of three-dimensional crystals. Protein Sci *18*, 2629-2635.

Ducat, D.C., and Silver, P.A. (2012). Improving Carbon Fixation Pathways. Current opinion in chemical biology *16*, 337-344.

Dvornyk, V., Vinogradova, O., and Nevo, E. (2003). Origin and evolution of circadian clock genes in prokaryotes. Proc Natl Acad Sci U S A *100*, 2495-2500.

Eisenhut, M., Aguirre von Wobeser, E., Jonas, L., Schubert, H., Ibelings, B.W., Bauwe, H., Matthijs, H.C., and Hagemann, M. (2007). Long-term response toward inorganic carbon limitation in wild type and glycolate turnover mutants of the cyanobacterium Synechocystis sp. strain PCC 6803. Plant Physiol *144*, 1946-1959.

Eisenhut, M., Ruth, W., Haimovich, M., Bauwe, H., Kaplan, A., and Hagemann, M. (2008). The photorespiratory glycolate metabolism is essential for cyanobacteria and might have been conveyed endosymbiontically to plants. Proc Natl Acad Sci U S A *105*, 17199-17204.

Eller, B.M., and Ferrari, S. (1997). Water use efficiency of two succulents with contrasting CO2 fixation pathways. Plant, Cell & Environment *20*, 93-100.

Emlyn-Jones, D., Woodger, F.J., Price, G.D., and Whitney, S.M. (2006). RbcX can function as a rubisco chaperonin, but is non-essential in Synechococcus PCC7942. Plant Cell Physiol *47*, 1630-1640.

Erbilgin, O., McDonald, K.L., and Kerfeld, C.A. (2014). Characterization of a Planctomycetal Organelle: a Novel Bacterial Microcompartment for the Aerobic Degradation of Plant Saccharides. Applied and environmental microbiology *80*, 2193-2205.

Espinosa, J., Boyd, J.S., Cantos, R., Salinas, P., Golden, S.S., and Contreras, A. (2015). Cross-talk and regulatory interactions between the essential response regulator RpaB and cyanobacterial circadian clock output. Proceedings of the National Academy of Sciences *112*, 2198-2203.

Fan, C., Cheng, S., Sinha, S., and Bobik, T.A. (2012). Interactions between the termini of lumen enzymes and shell proteins mediate enzyme encapsulation into bacterial microcompartments. Proc Natl Acad Sci U S A *109*, 14995-15000.

Fang, Y., Huang, F., Faulkner, M., Jiang, Q., Dykes, G.F., Yang, M., and Liu, L.N. (2018). Engineering and Modulating Functional Cyanobacterial CO2-Fixing Organelles. Front Plant Sci *9*, 739.

Faulkner, M., Rodriguez-Ramos, J., Dykes, G.F., Owen, S.V., Casella, S., Simpson, D.M., Beynon, R.J., and Liu, L.-N. (2017). Direct characterization of the native structure and mechanics of cyanobacterial carboxysomes. Nanoscale *9*, 10662–10673.

Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J.F., Dougherty, B.A., Merrick, J.M., *et al.* (1995). Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. Science *269*, 496-512.

Flombaum, P., Gallegos, J.L., Gordillo, R.A., Rincon, J., Zabala, L.L., Jiao, N., Karl, D.M., Li, W.K., Lomas, M.W., Veneziano, D., *et al.* (2013). Present and future global distributions of the marine Cyanobacteria Prochlorococcus and Synechococcus. Proc Natl Acad Sci U S A *110*, 9824-9829.

Fraser, C.M., Gocayne, J.D., White, O., Adams, M.D., Clayton, R.A., Fleischmann, R.D., Bult, C.J., Kerlavage, A.R., Sutton, G., Kelley, J.M., *et al.* (1995). The minimal gene complement of Mycoplasma genitalium. Science *270*, 397-403.

Frey, R., Mantri, S., Rocca, M., and Hilvert, D. (2016). Bottom-up Construction of a Primordial Carboxysome Mimic. J Am Chem Soc *138*, 10072-10075.

Froger, A., and Hall, J.E. (2007). Transformation of Plasmid DNA into E. coli Using the Heat Shock Method. JoVE, e253.

Fujita, Y., Murakami, A., and Ohki, K. (1987). Regulation of Photosystem Composition in the Cyanobacterial Photosynthetic System: the Regulation Occurs in Response to the Redox State of the Electron Pool Located between the Two Photosystems. Plant and Cell Physiology *28*, 283-292.

Furbank, R.T., Jenkins, C.L.D., and Hatch, M.D. (1989). CO(2) Concentrating Mechanism of C(4) Photosynthesis: Permeability of Isolated Bundle Sheath Cells to Inorganic Carbon. Plant Physiology *91*, 1364-1371.

Gabaldón, T., and Pittis, A.A. (2015). Origin and evolution of metabolic sub-cellular compartmentalization in eukaryotes. Biochimie *119*, 262-268.

Galmes, J., Kapralov, M.V., Andralojc, P.J., Conesa, M.A., Keys, A.J., Parry, M.A., and Flexas, J. (2014). Expanding knowledge of the Rubisco kinetics variability in plant species: environmental and evolutionary trends. Plant Cell Environ *37*, 1989-2001.

Garcia-Alles, L.F., Lesniewska, E., Root, K., Aubry, N., Pocholle, N., Mendoza, C.I., Bourillot, E., Barylyuk, K., Pompon, D., Zenobi, R., *et al.* (2017). Spontaneous non-canonical assembly of CcmK hexameric components from β-carboxysome shells of cyanobacteria. PLoS ONE *12*, e0185109.

Gill, R.T., Katsoulakis, E., Schmitt, W., Taroncher-Oldenburg, G., Misra, J., and Stephanopoulos, G. (2002). Genome-wide dynamic transcriptional profiling of the light-to-dark transition in Synechocystis sp. strain PCC 6803. J Bacteriol *184*, 3671-3681.

Goda, K., Kondo, T., and Oyama, T. (2014). Effects of adenylates on the circadian interaction of KaiB with the KaiC complex in the reconstituted cyanobacterial Kai protein oscillator. Biosci Biotechnol Biochem *78*, 1833-1838.

Golden, S.S. (2003). Timekeeping in bacteria: the cyanobacterial circadian clock. Current opinion in microbiology *6*, 535-540.

Golden, S.S., and Sherman, L.A. (1984). Optimal conditions for genetic transformation of the cyanobacterium Anacystis nidulans R2. J Bacteriol *158*, 36-42.

Gonzalez-Esquer, C.R., Shubitowski, T.B., and Kerfeld, C.A. (2015). Streamlined Construction of the Cyanobacterial CO(2)-Fixing Organelle via Protein Domain Fusions for Use in Plant Synthetic Biology. The Plant cell *27*, 2637-2644.

Gowik, U., and Westhoff, P. (2011). The path from C3 to C4 photosynthesis. Plant Physiol *155*, 56-63.
Griffiths, H. (2006). Designs on Rubisco. Nature *441*, 940.
Guerreiro, A.C.L., Benevento, M., Lehmann, R., van Breukelen, B., Post, H., Giansanti, P., Maarten Altelaar, A.F., Axmann, I.M., and Heck, A.J.R. (2014). Daily Rhythms in the Cyanobacterium Synechococcus elongatus Probed by High-resolution Mass Spectrometry–based Proteomics Reveals a Small Defined Set of Cyclic Proteins. Molecular & cellular proteomics : MCP *13*, 2042-2055.
Gutu, A., and O'Shea, Erin K. (2013). Two Antagonistic Clock-Regulated Histidine Kinases Time the Activation of Circadian Gene Expression. Molecular cell *50*, 288-294.
Hagemann, M., Fernie, A.R., Espie, G.S., Kern, R., Eisenhut, M., Reumann, S., Bauwe, H., and Weber, A.P. (2013). Evolution of the biochemistry of the photorespiratory C2 cycle. Plant biology (Stuttgart, Germany) *15*, 639-647.
Hagen, A., Sutter, M., Sloan, N., and Kerfeld, C.A. (2018). Programmed loading and rapid purification of engineered bacterial microcompartment shells. Nat Commun *9*, 2881.
Hanson, M.R., Lin, M.T., Carmo-Silva, A.E., and Parry, M.A. (2016). Towards engineering carboxysomes into C3 plants. Plant J *87*, 38-50.
Harano, K., Ishida, H., Kittaka, R., Kojima, K., Inoue, N., Tsukamoto, M., Satoh, R., Himeno, M., Iwaki, T., and Wadano, A. (2003). Regulation of the expression of ribulose-1,5-bisphosphate carboxylase/oxygenase (EC 4.1.1.39) in a cyanobacterium, Synechococcus PCC7942. Photosynth Res *78*, 59-65.
Harris, G.C., and Köniğer, M. (1997). The 'high' concentrations of enzymes within the chloroplast. Photosynthesis research *54*, 5-23.
Havemann, G.D., Sampson, E.M., and Bobik, T.A. (2002). PduA is a shell protein of polyhedral organelles involved in coenzyme B(12)-dependent degradation of 1,2-propanediol in Salmonella enterica serovar typhimurium LT2. J Bacteriol *184*, 1253-1261.
Heinhorst, S., Cannon, G.C., and Shively, J.M. (2014). Carboxysomes and Their Structural Organization in Prokaryotes. In Nanomicrobiology: Physiological and Environmental Characteristics, L.L. Barton, D.A. Bazylinski, and H. Xu, eds. (New York, NY: Springer New York), pp. 75-101.
Hihara, Y., Kamei, A., Kanehisa, M., Kaplan, A., and Ikeuchi, M. (2001). DNA Microarray Analysis of Cyanobacterial Gene Expression during Acclimation to High Light. The Plant cell *13*, 793-806.
Hoffman, R. (2002). Green fluorescent protein imaging of tumour growth, metastasis, and angiogenesis in mouse models. The Lancet Oncology *3*, 546-556.
Howitt, C.A., Cooley, J.W., Wiskich, J.T., and Vermaas, W.F. (2001). A strain of Synechocystis sp. PCC 6803 without photosynthetic oxygen evolution and respiratory oxygen consumption: implications for the study of cyclic photosynthetic electron transport. Planta *214*, 46-56.
Huang, L., McCluskey, M.P., Ni, H., and LaRossa, R.A. (2002). Global gene expression profiles of the cyanobacterium Synechocystis sp. strain PCC 6803 in response to irradiation with UV-B and white light. J Bacteriol *184*, 6845-6858.
Huang, X., Holden, H.M., and Raushel, F.M. (2001). Channeling of substrates and intermediates in enzyme-catalyzed reactions. Annual review of biochemistry *70*, 149-180.
Huff, J. (2015). The Airyscan detector from ZEISS: confocal imaging with improved signal-to-noise ratio and super-resolution. Nature Methods *12*, 1205.
Iancu, C.V., Ding, H.J., Morris, D.M., Dias, D.P., Gonzales, A.D., Martino, A., and Jensen, G.J. (2007). The structure of isolated Synechococcus strain WH8102 carboxysomes as revealed by electron cryotomography. J Mol Biol *372*, 764-773.
Iancu, C.V., Morris, D.M., Dou, Z., Heinhorst, S., Cannon, G.C., and Jensen, G.J. (2010). Organization, structure, and assembly of alpha-carboxysomes determined by electron cryotomography of intact cells. J Mol Biol *396*, 105-117.
Ishiura, M., Kutsuna, S., Aoki, S., Iwasaki, H., Andersson, C.R., Tanabe, A., Golden, S.S., Johnson, C.H., and Kondo, T. (1998). Expression of a gene cluster kaiABC as a circadian feedback process in cyanobacteria. Science *281*, 1519-1523.
Ito, H., Mutsuda, M., Murayama, Y., Tomita, J., Hosokawa, N., Terauchi, K., Sugita, C., Sugita, M., Kondo, T., and Iwasaki, H. (2009). Cyanobacterial daily life with Kai-based circadian and

diurnal genome-wide transcriptional control in Synechococcus elongatus. Proc Natl Acad Sci U S A *106*, 14168-14173.

Ivleva, N.B., Bramlett, M.R., Lindahl, P.A., and Golden, S.S. (2005). LdpA: a component of the circadian clock senses redox state of the cell. The EMBO journal *24*, 1202-1210.

Ivleva, N.B., Gao, T., LiWang, A.C., and Golden, S.S. (2006). Quinone sensing by the circadian input kinase of the cyanobacterial circadian clock. Proceedings of the National Academy of Sciences *103*, 17468-17473.

Iwasaki, H., Williams, S.B., Kitayama, Y., Ishiura, M., Golden, S.S., and Kondo, T. (2000). A KaiC-Interacting Sensory Histidine Kinase, SasA, Necessary to Sustain Robust Circadian Oscillation in Cyanobacteria. Cell *101*, 223-233.

Jain, I.H., Vijayan, V., and O'Shea, E.K. (2012). Spatial ordering of chromosomes enhances the fidelity of chromosome partitioning in cyanobacteria. Proc Natl Acad Sci U S A *109*, 13638-13643.

Jakobson, C.M., Slininger Lee, M.F., and Tullman-Ercek, D. (2017). De novo design of signal sequences to localize cargo to the 1,2-propanediol utilization microcompartment. Protein Sci *26*, 1086-1092.

Jiang, W., and Tang, L. (2017). Atomic cryo-EM structures of viruses. Current opinion in structural biology *46*, 122-129.

Johnson, C.H., Zhao, C., Xu, Y., and Mori, T. (2017). Timing the day: what makes bacterial clocks tick? Nat Rev Microbiol *15*, 232-242.

Keeley, J., xa, E, Rundel, P., xa, and W (2003). Evolution of CAM and C₄ Carbon‐Concentrating Mechanisms. International Journal of Plant Sciences *164*, S55-S77.

Kellogg, E.A. (2013). C4 photosynthesis. Current Biology *23*, R594-R599.

Kerfeld, C.A., Aussignargues, C., Zarzycki, J., Cai, F., and Sutter, M. (2018). Bacterial microcompartments. Nature Reviews Microbiology *16*, 277.

Kerfeld, C.A., and Erbilgin, O. (2015). Bacterial microcompartments and the modular construction of microbial metabolism. Trends in microbiology *23*, 22-34.

Kerfeld, C.A., Heinhorst, S., and Cannon, G.C. (2010). Bacterial microcompartments. Annu Rev Microbiol *64*, 391-408.

Kerfeld, C.A., and Melnicki, M.R. (2016). Assembly, function and evolution of cyanobacterial carboxysomes. Curr Opin Plant Biol *31*, 66-75.

Kerfeld, C.A., Sawaya, M.R., Tanaka, S., Nguyen, C.V., Phillips, M., Beeby, M., and Yeates, T.O. (2005). Protein structures forming the shell of primitive bacterial organelles. Science *309*, 936-938.

Kim, H.S., Cho, H.R., Choi, S.H., Woo, J.S., and Moon, W.K. (2010). In vivo imaging of tumor transduced with bimodal lentiviral vector encoding human ferritin and green fluorescent protein on a 1.5T clinical magnetic resonance scanner. Cancer research *70*, 7315-7324.

Klein, M.G., Zwart, P., Bagby, S.C., Cai, F., Chisholm, S.W., Heinhorst, S., Cannon, G.C., and Kerfeld, C.A. (2009). Identification and structural analysis of a novel carboxysome shell protein with implications for metabolite transport. J Mol Biol *392*, 319-333.

Kobayashi, M., Ishizuka, T., Katayama, M., Kanehisa, M., Bhattacharyya-Pakrasi, M., Pakrasi, H.B., and Ikeuchi, M. (2004). Response to oxidative stress involves a novel peroxiredoxin gene in the unicellular cyanobacterium Synechocystis sp. PCC 6803. Plant & cell physiology *45*, 290-299.

Kofoid, E., Rappleye, C., Stojiljkovic, I., and Roth, J. (1999). The 17-Gene Ethanolamine (eut) Operon of Salmonella typhimurium Encodes Five Homologues of Carboxysome Shell Proteins. Journal of Bacteriology *181*, 5317-5329.

Kolesinski, P., Belusiak, I., Czarnocki-Cieciura, M., and Szczepaniak, A. (2014). Rubisco Accumulation Factor 1 from Thermosynechococcus elongatus participates in the final stages of ribulose-1,5-bisphosphate carboxylase/oxygenase assembly in Escherichia coli cells and in vitro. The FEBS Journal *281*, 3920-3932.

Komenda, J., and Barber, J. (1995). Comparison of psbO and psbH deletion mutants of Synechocystis PCC 6803 indicates that degradation of D1 protein is regulated by the QB site and dependent on protein synthesis. Biochemistry *34*, 9625-9631.

Kraus, F., Miron, E., Demmerle, J., Chitiashvili, T., Budco, A., Alle, Q., Matsuda, A., Leonhardt, H., Schermelleh, L., and Markaki, Y. (2017). Quantitative 3D structured illumination microscopy of nuclear structures. Nature Protocols *12*, 1011.

Kuznetsov, Y.G., and McPherson, A. (2011). Atomic Force Microscopy in Imaging of Viruses and Virus-Infected Cells. Microbiology and Molecular Biology Reviews : MMBR *75*, 268-285.

Laloux, G., and Jacobs-Wagner, C. (2014). How do bacteria localize proteins to the cell pole? Journal of cell science *127*, 11.

Larsson, A.M., Hasse, D., Valegard, K., and Andersson, I. (2017). Crystal structures of β-carboxysome shell protein CcmP: ligand binding correlates with the closed or open central pore. J Exp Bot *68*, 3857-3867.

Lassila, J.K., Bernstein, S.L., Kinney, J.N., Axen, S.D., and Kerfeld, C.A. (2014). Assembly of Robust Bacterial Microcompartment Shells Using Building Blocks from an Organelle of Unknown Function. Journal of Molecular Biology *426*, 2217-2228.

Lea-Smith, D.J., Bombelli, P., Vasudevan, R., and Howe, C.J. (2016). Photosynthetic, respiratory and extracellular electron transport pathways in cyanobacteria. Biochimica et Biophysica Acta (BBA) - Bioenergetics *1857*, 247-255.

Leake, M.C. (2018). Transcription factors in eukaryotic cells can functionally regulate gene expression by acting in oligomeric assemblies formed from an intrinsically disordered protein phase transition enabled by molecular crowding. Transcription.

Leake, M.C., Chandler, J.H., Wadhams, G.H., Bai, F., Berry, R.M., and Armitage, J.P. (2006). Stoichiometry and turnover in single, functioning membrane protein complexes. Nature *443*, 355-358.

Leake, M.C., Greene, N.P., Godun, R.M., Granjon, T., Buchanan, G., Chen, S., Berry, R.M., Palmer, T., and Berks, B.C. (2008). Variable stoichiometry of the TatA component of the twin-arginine protein transport system observed by in vivo single-molecule imaging. Proc Natl Acad Sci U S A *105*, 15376-15381.

Leake, M.C., Wilson, D., Bullard, B., and Simmons, R.M. (2003). The elasticity of single kettin molecules using a two-bead laser-tweezers assay. FEBS Lett *535*, 55-60.

Leake, M.C., Wilson, D., Gautel, M., and Simmons, R.M. (2004). The elasticity of single titin molecules using a two-bead optical tweezers assay. Biophysical journal *87*, 1112-1135.

Lee, K.K., and Gui, L. (2016). Dissecting Virus Infectious Cycles by Cryo-Electron Microscopy. PLOS Pathogens *12*, e1005625.

Li, Q., Yafal, A.G., Lee, Y.M., Hogle, J., and Chow, M. (1994). Poliovirus neutralization by antibodies to internal epitopes of VP4 and VP1 results from reversible exposure of these sequences at physiological temperature. J Virol *68*, 3965-3970.

Liberton, M., Austin, J.R., 2nd, Berg, R.H., and Pakrasi, H.B. (2011). Unique thylakoid membrane architecture of a unicellular N_2-fixing cyanobacterium revealed by electron tomography. Plant Physiol *155*, 1656-1666.

Lin, M.T., Occhialini, A., Andralojc, P.J., Devonshire, J., Hines, K.M., Parry, M.A., and Hanson, M.R. (2014a). beta-Carboxysomal proteins assemble into highly organized structures in Nicotiana chloroplasts. Plant J *79*, 1-12.

Lin, M.T., Occhialini, A., Andralojc, P.J., Parry, M.A., and Hanson, M.R. (2014b). A faster Rubisco with potential to increase photosynthesis in crops. Nature *513*, 547-550.

Liu, L.N. (2016). Distribution and dynamics of electron transport complexes in cyanobacterial thylakoid membranes. Biochimica et biophysica acta *1857*, 256-265.

Liu, L.N., Bryan, S.J., Huang, F., Yu, J., Nixon, P.J., Rich, P.R., and Mullineaux, C.W. (2012). Control of electron transport routes through redox-regulated redistribution of respiratory complexes. Proc Natl Acad Sci U S A *109*, 11431-11436.

Liu, L.N., and Scheuring, S. (2013). Investigation of photosynthetic membrane structure using atomic force microscopy. Trends in plant science *18*, 277-286.

Liu, Z., Taub, C.C., and McClung, C.R. (1996). Identification of an Arabidopsis thaliana ribulose-1,5-bisphosphate carboxylase/oxygenase activase (RCA) minimal promoter regulated by light and the circadian clock. Plant Physiol *112*, 43-51.

Long, B.M., Badger, M.R., Whitney, S.M., and Price, G.D. (2007). Analysis of carboxysomes from Synechococcus PCC7942 reveals multiple Rubisco complexes with carboxysomal proteins CcmM and CcaA. J Biol Chem *282*, 29323-29335.

Long, B.M., Hee, W.Y., Sharwood, R.E., Rae, B.D., Kaines, S., Lim, Y.-L., Nguyen, N.D., Massey, B., Bala, S., von Caemmerer, S., *et al.* (2018). Carboxysome encapsulation of the CO2-fixing enzyme Rubisco in tobacco chloroplasts. Nature Communications *9*, 3570.

Long, B.M., Price, G.D., and Badger, M.R. (2005). Proteomic assessment of an established technique for carboxysome enrichment from Synechococcus PCC7942. Canadian Journal of Botany *83*, 746-757.

Long, B.M., Rae, B.D., Badger, M.R., and Price, G.D. (2011). Over-expression of the beta-carboxysomal CcmM protein in Synechococcus PCC7942 reveals a tight co-regulation of carboxysomal carbonic anhydrase (CcaA) and M58 content. Photosynth Res *109*, 33-45.

Long, B.M., Tucker, L., Badger, M.R., and Price, G.D. (2010). Functional cyanobacterial β-carboxysomes have an absolute requirement for both long and short forms of the CcmM protein. Plant Physiol *153*, 285-293.

Loza-Correa, M., Gomez-Valero, L., and Buchrieser, C. (2010). Circadian Clock Proteins in Prokaryotes: Hidden Rhythms? Frontiers in Microbiology *1*, 130.

Lund, V.A., Wacnik, K., Turner, R.D., Cotterell, B.E., Walther, C.G., Fenn, S.J., Grein, F., Wollman, A.J., Leake, M.C., Olivier, N., *et al.* (2018). Molecular coordination of *Staphylococcus aureus* cell division. Elife *7*.

MacCready, J.S., Hakim, P., Young, E.J., Hu, L., Liu, J., Osteryoung, K.W., Vecchiarelli, A.G., and Ducat, D.C. (2018). Protein Gradients on the Nucleoid Position the Carbon-fixing Organelles of Cyanobacteria. bioRxiv.

Mackey, S.R., Ditty, J.L., Clerico, E.M., and Golden, S.S. (2007). Detection of rhythmic bioluminescence from luciferase reporters in cyanobacteria. Methods Mol Biol *362*, 115-129.

Maeda, S., Badger, M.R., and Price, G.D. (2002). Novel gene products associated with NdhD3/D4-containing NDH-1 complexes are involved in photosynthetic CO2 hydration in the cyanobacterium, Synechococcus sp. PCC7942. Mol Microbiol *43*, 425-435.

Mahinthichaichan, P., Morris, D.M., Wang, Y., Jensen, G.J., and Tajkhorshid, E. (2018). Selective Permeability of Carboxysome Shell Pores to Anionic Molecules. bioRxiv.

Mangan, N., and Brenner, M. (2014). Systems analysis of the CO2 concentrating mechanism in cyanobacteria. Elife, e02043.

Marchetti, M., Wuite, G.J.L., and Roos, W.H. (2016). Atomic force microscopy observation and characterization of single virions and virus-like particles by nano-indentation. Current Opinion in Virology *18*, 82-88.

Marco, E., Martinez, I., Ronen-Tarazi, M., Orus, M.I., and Kaplan, A. (1994). Inactivation of ccmO in Synechococcus sp. Strain PCC 7942 Results in a Mutant Requiring High Levels of CO(2). Appl Environ Microbiol *60*, 1018-1020.

Masojídek, J., Grobbelaar, J.U., Pechar, L., and KoblížEk, M. (2001). Photosystem II Electron Transport Rates and Oxygen Production in Natural Waterblooms of Freshwater Cyanobacteria During a Diel Cycle. Journal of Plankton Research *23*, 57-66.

Mayer, M.J., Juodeikis, R., Brown, I.R., Frank, S., Palmer, D.J., Deery, E., Beal, D.M., Xue, W.-F., and Warren, M.J. (2016). Effect of bio-engineering on size, shape, composition and rigidity of bacterial microcompartments. Scientific Reports *6*, 36899.

McGinn, P.J., Price, G.D., Maleszka, R., and Badger, M.R. (2003). Inorganic Carbon Limitation and Light Control the Expression of Transcripts Related to the CO(2)-Concentrating Mechanism in the Cyanobacterium Synechocystis sp. Strain PCC6803. Plant Physiology *132*, 218-229.

McGinn., M.P., D., P.G., and R., B.M. (2004). High light enhances the expression of low-CO2-inducible transcripts involved in the CO2-concentrating mechanism in Synechocystis sp. PCC6803. Plant, Cell & Environment *27*, 615-626.

McGurn, L.D., Moazami-Goudarzi, M., White, S.A., Suwal, T., Brar, B., Tang, J.Q., Espie, G.S., and Kimber, M.S. (2016). The structure, kinetics and interactions of the beta-carboxysomal beta-carbonic anhydrase, CcaA. Biochem J *473*, 4559-4572.

McKay, R.M.L., Gibbs, S.P., and Espie, G.S. (1993). Effect of dissolved inorganic carbon on the expression of carboxysomes, localization of Rubisco and the mode of inorganic carbon

transport in cells of the cyanobacterium Synechococcus UTEX 625. Archives of Microbiology *159*, 21-29.

McNevin, D., von Caemmerer, S., and Farquhar, G. (2006). Determining RuBisCO activation kinetics and other rate and equilibrium constants by simultaneous multiple non-linear regression of a kinetic model. J Exp Bot *57*, 3883-3900.

Menon, B.B., Dou, Z., Heinhorst, S., Shively, J.M., and Cannon, G.C. (2008). Halothiobacillus neapolitanus carboxysomes sequester heterologous and chimeric RubisCO species. PLoS One *3*, e3570.

Menon, B.B., Heinhorst, S., Shively, J.M., and Cannon, G.C. (2010). The carboxysome shell is permeable to protons. J Bacteriol *192*, 5881-5886.

Michelet, L., Zaffagnini, M., Morisse, S., Sparla, F., Pérez-Pérez, M.E., Francia, F., Danon, A., Marchand, C.H., Fermani, S., Trost, P., *et al.* (2013). Redox regulation of the Calvin–Benson cycle: something old, something new. Frontiers in Plant Science *4*, 470.

Miller, H., Cosgrove, J., Wollman, A.J.M., Taylor, E., Zhou, Z., O'Toole, P.J., Coles, M.C., and Leake, M.C. (2018). High-speed single-molecule tracking of CXCL13 in the B-follicle. Front Immunol *9*, 1073.

Miller, H., Zhou, Z., Wollman, A.J., and Leake, M.C. (2015). Superresolution imaging of single DNA molecules using stochastic photoblinking of minor groove and intercalating dyes. Methods *88*, 81-88.

Montgomery, B.L., Lechno-Yossef, S., and Kerfeld, C.A. (2016). Interrelated modules in cyanobacterial photosynthesis: the carbon-concentrating mechanism, photorespiration, and light perception. J Exp Bot *67*, 2931-2940.

Mori, T., Binder, B., and Johnson, C.H. (1996). Circadian gating of cell division in cyanobacteria growing with average doubling times of less than 24 hours. Proceedings of the National Academy of Sciences *93*, 10183-10188.

Mullineaux, C.W. (2001). How do cyanobacteria sense and respond to light? Mol Microbiol *41*, 965-971.

Mullineaux, C.W., and Allen, J.F. (1990). State 1-State 2 transitions in the cyanobacterium Synechococcus 6301 are controlled by the redox state of electron carriers between Photosystems I and II. Photosynth Res *23*, 297-311.

Muzzey, D., and van Oudenaarden, A. (2009). Quantitative time-lapse fluorescence microscopy in single cells. Annual review of cell and developmental biology *25*, 301-327.

Nakajima, M., Imai, K., Ito, H., Nishiwaki, T., Murayama, Y., Iwasaki, H., Oyama, T., and Kondo, T. (2005). Reconstitution of circadian oscillation of cyanobacterial KaiC phosphorylation in vitro. Science *308*, 414-415.

Nassoury, N., Fritz, L., and Morse, D. (2001). Circadian Changes in Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase Distribution Inside Individual Chloroplasts Can Account for the Rhythm in Dinoflagellate Carbon Fixation. The Plant Cell *13*, 923.

Niederhuber, M.J., Lambert, T.J., Yapp, C., Silver, P.A., and Polka, J.K. (2017). Superresolution microscopy of the β-carboxysome reveals a homogeneous matrix. Molecular Biology of the Cell *28*, 2734-2745.

Nishiwaki, T., Satomi, Y., Nakajima, M., Lee, C., Kiyohara, R., Kageyama, H., Kitayama, Y., Temamoto, M., Yamaguchi, A., Hijikata, A., *et al.* (2004). Role of KaiC phosphorylation in the circadian clock system of Synechococcus elongatus PCC 7942. Proc Natl Acad Sci U S A *101*, 13927-13932.

Nogales, J., Gudmundsson, S., Knight, E.M., Palsson, B.O., and Thiele, I. (2012). Detailing the optimality of photosynthesis in cyanobacteria through systems biology analysis. Proceedings of the National Academy of Sciences of the United States of America *109*, 2678-2683.

Novak, D., Kucharova, A., Ovecka, M., Komis, G., and Samaj, J. (2015). Developmental Nuclear Localization and Quantification of GFP-Tagged EB1c in Arabidopsis Root Using Light-Sheet Microscopy. Front Plant Sci *6*, 1187.

Occhialini, A., Lin, M.T., Andralojc, P.J., Hanson, M.R., and Parry, M.A. (2016). Transgenic tobacco plants with improved cyanobacterial Rubisco expression but no extra assembly factors grow at near wild-type rates if provided with elevated CO2. Plant J *85*, 148-160.

Ohbayashi, R., Watanabe, S., Kanesaki, Y., Narikawa, R., Chibazakura, T., Ikeuchi, M., and Yoshikawa, H. (2013). DNA replication depends on photosynthetic electron transport in cyanobacteria. FEMS Microbiol Lett *344*, 138-144.

Omata, T., Price, G.D., Badger, M.R., Okamura, M., Gohta, S., and Ogawa, T. (1999). Identification of an ATP-binding cassette transporter involved in bicarbonate uptake in the cyanobacterium Synechococcus sp. strain PCC 7942. Proceedings of the National Academy of Sciences *96*, 13571-13576.

Onizuka, T., Endo, S., Akiyama, H., Kanai, S., Hirano, M., Yokota, A., Tanaka, S., and Miyasaka, H. (2004). The rbcX gene product promotes the production and assembly of ribulose-1,5-bisphosphate carboxylase/oxygenase of Synechococcus sp. PCC7002 in Escherichia coli. Plant Cell Physiol *45*, 1390-1395.

Osanai, T., Imamura, S., Asayama, M., Shirai, M., Suzuki, I., Murata, N., and Tanaka, K. (2006). Nitrogen induction of sugar catabolic gene expression in Synechocystis sp. PCC 6803. DNA research : an international journal for rapid publication of reports on genes and genomes *13*, 185-195.

Paddock, M.L., Boyd, J.S., Adin, D.M., and Golden, S.S. (2013). Active output state of the Synechococcus Kai circadian oscillator. Proceedings of the National Academy of Sciences *110*, E3849-E3857.

Park, J., Chun, S., Bobik, T.A., Houk, K.N., and Yeates, T.O. (2017). Molecular Dynamics Simulations of Selective Metabolite Transport across the Propanediol Bacterial Microcompartment Shell. The journal of physical chemistry B *121*, 8149-8154.

Parry, Bradley R., Surovtsev, Ivan V., Cabeen, Matthew T., O'Hern, Corey S., Dufresne, Eric R., and Jacobs-Wagner, C. (2014). The Bacterial Cytoplasm Has Glass-like Properties and Is Fluidized by Metabolic Activity. Cell *156*, 183-194.

Parsons, J.B., Frank, S., Bhella, D., Liang, M., Prentice, M.B., Mulvihill, D.P., and Warren, M.J. (2010). Synthesis of Empty Bacterial Microcompartments, Directed Organelle Protein Incorporation, and Evidence of Filament-Associated Organelle Movement. Molecular Cell *38*, 305-315.

Pena, K.L., Castel, S.E., de Araujo, C., Espie, G.S., and Kimber, M.S. (2010). Structural basis of the oxidative activation of the carboxysomal gamma-carbonic anhydrase, CcmM. Proc Natl Acad Sci U S A *107*, 2455-2460.

Perlmutter, J.D., Mohajerani, F., and Hagan, M.F. (2016). Many-molecule encapsulation by an icosahedral shell. eLife *5*, e14078.

Pernas, L., and Scorrano, L. (2016). Mito-Morphosis: Mitochondrial Fusion, Fission, and Cristae Remodeling as Key Mediators of Cellular Function. Annual review of physiology *78*, 505-531.

Plank, M., Wadhams, G.H., and Leake, M.C. (2009). Millisecond timescale slimfield imaging and automated quantification of single fluorescent protein molecules for use in probing complex biological processes. Integr Biol (Camb) *1*, 602-612.

Price, G.D., and Badger, M.R. (1989). Isolation and Characterization of High CO(2)-Requiring-Mutants of the Cyanobacterium Synechococcus PCC7942 : Two Phenotypes that Accumulate Inorganic Carbon but Are Apparently Unable to Generate CO(2) within the Carboxysome. Plant Physiol *91*, 514-525.

Price, G.D., and Badger, M.R. (1991). Evidence for the role of carboxysomes in the cyanobacterial CO_2-concentrating mechanism. Can J Bot *69*, 963-973.

Price, G.D., Badger, M.R., Woodger, F.J., and Long, B.M. (2008). Advances in understanding the cyanobacterial CO2-concentrating-mechanism (CCM): functional components, Ci transporters, diversity, genetic regulation and prospects for engineering into plants. Journal of Experimental Botany *59*, 1441-1461.

Price, G.D., Coleman, J.R., and Badger, M.R. (1992). Association of Carbonic Anhydrase Activity with Carboxysomes Isolated from the Cyanobacterium Synechococcus PCC7942. Plant Physiol *100*, 784-793.

Price, G.D., Sültemeyer, D., Klughammer, B., Ludwig, M., and Badger, M.R. (1998). The functioning of the CO2 concentrating mechanism in several cyanobacterial strains: a review of general physiological characteristics, genes, proteins, and recent advances. Canadian Journal of Botany *76*, 973-1002.

Price, G.D., Woodger, F.J., Badger, M.R., Howitt, S.M., and Tucker, L. (2004). Identification of a SulP-type bicarbonate transporter in marine cyanobacteria. Proc Natl Acad Sci U S A *101*, 18228-18233.

Rae, B.D., Long, B.M., Badger, M.R., and Price, G.D. (2012). Structural determinants of the outer shell of beta-carboxysomes in Synechococcus elongatus PCC 7942: roles for CcmK2, K3-K4, CcmO, and CcmL. PLoS One *7*, e43871.

Rae, B.D., Long, B.M., Badger, M.R., and Price, G.D. (2013). Functions, compositions, and evolution of the two types of carboxysomes: polyhedral microcompartments that facilitate CO2 fixation in cyanobacteria and some proteobacteria. Microbiol Mol Biol Rev *77*, 357-379.

Reyes-Lamothe, R., Sherratt, D.J., and Leake, M.C. (2010). Stoichiometry and architecture of active DNA replication machinery in Escherichia coli. Science *328*, 498-501.

Ringgaard, S., van Zon, J., Howard, M., and Gerdes, K. (2009). Movement and equipositioning of plasmids by ParA filament disassembly. Proc Natl Acad Sci U S A *106*, 19369-19374.

Rippka, R., Deruelles, J., Waterbury, J.B., Herdman, M., and Stanier, R.Y. (1979). Generic Assignments, Strain Histories and Properties of Pure Cultures of Cyanobacteria. Microbiology *111*, 1-61.

Rodriguez-Ramos, J., Faulkner, M., and Liu, L. (2018). Nanoscale Visualization of Bacterial Microcompartments Using Atomic Force Microscopy, Vol 1814.

Rohnke, B.A., Singh, S.P., Pattanaik, B., and Montgomery, B.L. (2018). RcaE-Dependent Regulation of Carboxysome Structural Proteins Has a Central Role in Environmental Determination of Carboxysome Morphology and Abundance in Fremyella diplosiphon. mSphere *3*, e00617-00617.

Roivainen, M., Piirainen, L., Rysa, T., Narvanen, A., and Hovi, T. (1993). An immunodominant N-terminal region of VP1 protein of poliovirion that is buried in crystal structure can be exposed in solution. Virology *195*, 762-765.

Rotskoff, G.M., and Geissler, P.L. (2018). Robust nonequilibrium pathways to microcompartment assembly. Proceedings of the National Academy of Sciences.

Rudner, D.Z., and Losick, R. (2010). Protein subcellular localization in bacteria. Cold Spring Harbor perspectives in biology *2*, a000307.

Rust, M.J., Golden, S.S., and O'Shea, E.K. (2011). Light-driven changes in energy metabolism directly entrain the cyanobacterial circadian oscillator. Science *331*, 220-223.

Samborska, B., and Kimber, Matthew S. (2012). A Dodecameric CcmK2 Structure Suggests β-Carboxysomal Shell Facets Have a Double-Layered Organization. Structure *20*, 1353-1362.

Sargent, F., Davidson, F.A., Kelly, C.L., Binny, R., Christodoulides, N., Gibson, D., Johansson, E., Kozyrska, K., Lado, L.L., MacCallum, J., *et al.* (2013). A synthetic system for expression of components of a bacterial microcompartment. Microbiology *159*, 2427-2436.

Saschenbrecker, S., Bracher, A., Rao, K.V., Rao, B.V., Hartl, F.U., and Hayer-Hartl, M. (2007). Structure and function of RbcX, an assembly chaperone for hexadecameric Rubisco. Cell *129*, 1189-1200.

Satori, C.P., Henderson, M.M., Krautkramer, E.A., Kostal, V., Distefano, M.M., and Arriaga, E.A. (2013). Bioanalysis of eukaryotic organelles. Chemical reviews *113*, 2733-2811.

Savage, D.F., Afonso, B., Chen, A.H., and Silver, P.A. (2010). Spatially ordered dynamics of the bacterial carbon fixation machinery. Science *327*, 1258-1261.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., *et al.* (2012). Fiji: an open-source platform for biological-image analysis. Nat Meth *9*, 676-682.

Schmid, M.F., Paredes, A.M., Khant, H.A., Soyer, F., Aldrich, H.C., Chiu, W., and Shively, J.M. (2006). Structure of Halothiobacillus neapolitanus Carboxysomes by Cryo-Electron Tomography. Journal of molecular biology *364*, 526-535.

Schneider, G., Lindqvist, Y., and Branden, C.I. (1992). RUBISCO: structure and mechanism. Annual review of biophysics and biomolecular structure *21*, 119-143.

Schopf, J.W., and Packer, B.M. (1987). Early Archean (3.3-billion to 3.5-billion-year-old) microfossils from Warrawoona Group, Australia. Science *237*, 70-73.

Schuurmans, R.M., Schuurmans, J.M., Bekker, M., Kromkamp, J.C., Matthijs, H.C.P., and Hellingwerf, K.J. (2014). The Redox Potential of the Plastoquinone Pool of the

Cyanobacterium Synechocystis Species Strain PCC 6803 Is under Strict Homeostatic Control. Plant Physiology *165*, 463-475.
Schwarz, D., Nodop, A., Huge, J., Purfurst, S., Forchhammer, K., Michel, K.P., Bauwe, H., Kopka, J., and Hagemann, M. (2011). Metabolic and transcriptomic phenotyping of inorganic carbon acclimation in the Cyanobacterium Synechococcus elongatus PCC 7942. Plant Physiol *155*, 1640-1655.
Schwarz, R., and Grossman, A.R. (1998). A response regulator of cyanobacteria integrates diverse environmental signals and is critical for survival under extreme conditions. Proc Natl Acad Sci U S A *95*, 11008-11013.
Schwarz, R., Reinhold, L., and Kaplan, A. (1995). Low Activation State of Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase in Carboxysome-Defective Synechococcus Mutants. Plant Physiol *108*, 183-190.
Shapiro, L., McAdams, H.H., and Losick, R. (2009). Why and How Bacteria Localize Proteins. Science *326*, 1225.
Shashkova, S., Wollman, A.J., Hohmann, S., and Leake, M.C. (2018). Characterising Maturation of GFP and mCherry of Genomically Integrated Fusions in Saccharomyces cerevisiae. Bio-protocol *8*, e2710.
Shestakov, S.V., and Khyen, N.T. (1970). Evidence for genetic transformation in blue-green alga Anacystis nidulans. Molecular & general genetics : MGG *107*, 372-375.
Shibata, M., Katoh, H., Sonoda, M., Ohkawa, H., Shimoyama, M., Fukuzawa, H., Kaplan, A., and Ogawa, T. (2002). Genes essential to sodium-dependent bicarbonate transport in cyanobacteria: function and phylogenetic analysis. J Biol Chem *277*, 18658-18664.
Shibata, M., Ohkawa, H., Kaneko, T., Fukuzawa, H., Tabata, S., Kaplan, A., and Ogawa, T. (2001). Distinct constitutive and low-CO2-induced CO2 uptake systems in cyanobacteria: genes involved and their phylogenetic relationship with homologous genes in other organisms. Proc Natl Acad Sci U S A *98*, 11789-11794.
Shively, J.M., Ball, F., Brown, D.H., and Saunders, R.E. (1973). Functional organelles in prokaryotes: polyhedral inclusions (carboxysomes) of Thiobacillus neapolitanus. Science *182*, 584-586.
Sirohi, D., Chen, Z., Sun, L., Klose, T., Pierson, T.C., Rossmann, M.G., and Kuhn, R.J. (2016). The 3.8 Å resolution cryo-EM structure of Zika virus. Science.
Smith, R.M., and Williams, S.B. (2006). Circadian rhythms in gene transcription imparted by chromosome compaction in the cyanobacterium Synechococcus elongatus. Proceedings of the National Academy of Sciences *103*, 8564-8569.
So, A.K.-C., Cot, S.S.-W., and Espie, G.S. (2002). Characterization of the C-terminal extension of carboxysomal carbonic anhydrase from <emph type="2">Synechocystis</emph> sp. PCC6803. Functional Plant Biology *29*, 183-194.
So, A.K.C., Espie, G.S., Williams, E.B., Shively, J.M., Heinhorst, S., and Cannon, G.C. (2004). A Novel Evolutionary Lineage of Carbonic Anhydrase (ε Class) Is a Component of the Carboxysome Shell. Journal of Bacteriology *186*, 623-630.
Sommer, M., Cai, F., Melnicki, M., and Kerfeld, C.A. (2017). beta-Carboxysome bioinformatics: identification and evolution of new bacterial microcompartment protein gene classes and core locus constraints. J Exp Bot *68*, 3841-3855.
Spitzer, J., and Poolman, B. (2013). How crowded is the prokaryotic cytoplasm? FEBS Letters *587*, 2094-2098.
Spreitzer, R.J. (2003). Role of the small subunit in ribulose-1,5-bisphosphate carboxylase/oxygenase. Archives of biochemistry and biophysics *414*, 141-149.
Sun, Y., Casella, S., Fang, Y., Huang, F., Faulkner, M., Barrett, S., and Liu, L.N. (2016). Light modulates the biosynthesis and organization of cyanobacterial carbon fixation machinery through photosynthetic electron flow. Plant Physiol *171*, 530-541.
Sutter, M., Faulkner, M., Aussignargues, C., Paasch, B.C., Barrett, S., Kerfeld, C.A., and Liu, L.N. (2016). Visualization of Bacterial Microcompartment Facet Assembly Using High-Speed Atomic Force Microscopy. Nano Lett *16*, 1590-1595.
Sutter, M., Greber, B., Aussignargues, C., and Kerfeld, C.A. (2017). Assembly principles and structure of a 6.5-MDa bacterial microcompartment shell. Science *356*, 1293-1297.

Sutter, M., Roberts, E.W., Gonzalez, R.C., Bates, C., Dawoud, S., Landry, K., Cannon, G.C., Heinhorst, S., and Kerfeld, C.A. (2015). Structural Characterization of a Newly Identified Component of α-Carboxysomes: The AAA+ Domain Protein CsoCbbQ. Scientific Reports *5*, 16243.

Sutter, M., Wilson, S.C., Deutsch, S., and Kerfeld, C.A. (2013). Two new high-resolution crystal structures of carboxysome pentamer proteins reveal high structural conservation of CcmL orthologs among distantly related cyanobacterial species. Photosynth Res *118*, 9-16.

Swan, J.A., Golden, S., LiWang, A., and Partch, C.L. (2018). Structure, function, and mechanism of the core circadian clock in cyanobacteria. Journal of Biological Chemistry.

Takai, N., Nakajima, M., Oyama, T., Kito, R., Sugita, C., Sugita, M., Kondo, T., and Iwasaki, H. (2006). A KaiC-associating SasA–RpaA two-component regulatory system as a major circadian timing mediator in cyanobacteria. Proceedings of the National Academy of Sciences *103*, 12109-12114.

Tamoi, M., Miyazaki, T., Fukamizo, T., and Shigeoka, S. (2005). The Calvin cycle in cyanobacteria is regulated by CP12 via the NAD(H)/NADP(H) ratio under light/dark conditions. Plant J *42*, 504-513.

Tanaka, S., Kerfeld, C.A., Sawaya, M.R., Cai, F., Heinhorst, S., Cannon, G.C., and Yeates, T.O. (2008). Atomic-level models of the bacterial carboxysome shell. Science *319*, 1083-1086.

Tanaka, S., Sawaya, M.R., Kerfeld, C.A., and Yeates, T.O. (2007). Structure of the RuBisCO chaperone RbcX from Synechocystis sp. PCC6803. Acta Crystallogr D Biol Crystallogr *63*, 1109-1112.

Tanaka, S., Sawaya, M.R., Phillips, M., and Yeates, T.O. (2009). Insights from multiple structures of the shell proteins from the β-carboxysome. Protein Science *18*, 108-120.

Taniguchi, Y., Takai, N., Katayama, M., Kondo, T., and Oyama, T. (2010). Three major output pathways from the KaiABC-based oscillator cooperate to generate robust circadian kaiBC expression in cyanobacteria. Proceedings of the National Academy of Sciences *107*, 3263-3268.

Tchernov, D., Helman, Y., Keren, N., Luz, B., Ohad, I., Reinhold, L., Ogawa, T., and Kaplan, A. (2001). Passive entry of CO2 and its energy-dependent intracellular conversion to HCO3- in cyanobacteria are driven by a photosystem I-generated deltamuH+. J Biol Chem *276*, 23450-23455.

Tilzer, M.M. (1987). Light‐dependence of photosynthesis and growth in cyanobacteria: Implications for their dominance in eutrophic lakes. New Zealand Journal of Marine and Freshwater Research *21*, 401-412.

Ting, C.S., Hsieh, C., Sundararaman, S., Mannella, C., and Marko, M. (2007). Cryo-electron tomography reveals the comparative three-dimensional architecture of Prochlorococcus, a globally important marine cyanobacterium. J Bacteriol *189*, 4485-4493.

Ting, I.P. (1985). Crassulacean Acid Metabolism. Annual Review of Plant Physiology *36*, 595-622.

To, K.Y., Suen, D.F., and Chen, S.C. (1999). Molecular characterization of ribulose-1,5-bisphosphate carboxylase/oxygenase activase in rice leaves. Planta *209*, 66-76.

Trebst, A. (1980). [65] Inhibitors in electron flow: Tools for the functional and structural localization of carriers and energy conservation sites. In Methods in Enzymology, A. San Pietro, ed. (Academic Press), pp. 675-715.

Tsai, Y.-C.C., Lapina, M.C., Bhushan, S., and Mueller-Cajar, O. (2015). Identification and characterization of multiple rubisco activases in chemoautotrophic bacteria. Nature Communications *6*, 8883.

Tsai, Y., Sawaya, M.R., Cannon, G.C., Cai, F., Williams, E.B., Heinhorst, S., Kerfeld, C.A., and Yeates, T.O. (2007). Structural analysis of CsoS1A and the protein shell of the Halothiobacillus neapolitanus carboxysome. PLoS Biol *5*, e144.

Tsinoremas, N.F., Schaefer, M.R., and Golden, S.S. (1994). Blue and red light reversibly control psbA expression in the cyanobacterium Synechococcus sp. strain PCC 7942. J Biol Chem *269*, 16143-16147.

Ungerer, J., Lin, P.C., Chen, H.Y., and Pakrasi, H.B. (2018). Adjustments to Photosystem Stoichiometry and Electron Transfer Proteins Are Key to the Remarkably Fast Growth of the Cyanobacterium Synechococcus elongatus UTEX 2973. mBio *9*.

Vecchiarelli, A.G., Hwang, L.C., and Mizuuchi, K. (2013). Cell-free study of F plasmid partition provides evidence for cargo transport by a diffusion-ratchet mechanism. Proc Natl Acad Sci U S A *110*, E1390-1397.

Vermaas, W. (2001). Photosynthesis and Respiration in Cyanobacteria, Vol 1–7.

Verveer, P.J., Wouters, F.S., Reynolds, A.R., and Bastiaens, P.I. (2000). Quantitative imaging of lateral ErbB1 receptor signal propagation in the plasma membrane. Science *290*, 1567-1570.

Vijayan, V., Jain, I.H., and O'Shea, E.K. (2011). A high resolution map of a cyanobacterial transcriptome. Genome biology *12*, R47.

Vijayan, V., Zuzow, R., and O'Shea, E.K. (2009). Oscillations in supercoiling drive circadian gene expression in cyanobacteria. Proc Natl Acad Sci U S A *106*, 22564-22568.

Wang, H.L., Postier, B.L., and Burnap, R.L. (2004). Alterations in global patterns of gene expression in Synechocystis sp. PCC 6803 in response to inorganic carbon limitation and the inactivation of ndhR, a LysR family regulator. J Biol Chem *279*, 5739-5751.

Watson, G.M., and Tabita, F.R. (1996). Regulation, unique gene organization, and unusual primary structure of carbon fixation genes from a marine phycoerythrin-containing cyanobacterium. Plant molecular biology *32*, 1103-1115.

Welkie, D.G., Rubin, B.E., Chang, Y.G., Diamond, S., Rifkin, S.A., LiWang, A., and Golden, S.S. (2018). Genome-wide fitness assessment during diurnal growth reveals an expanded role of the cyanobacterial circadian clock protein KaiA. Proc Natl Acad Sci U S A.

Wheatley, N.M., Gidaniyan, S.D., Liu, Y., Cascio, D., and Yeates, T.O. (2013). Bacterial microcompartment shells of diverse functional types possess pentameric vertex proteins. Protein Sci *22*, 660-665.

Whitehead, L., Long, B.M., Price, G.D., and Badger, M.R. (2014). Comparing the in vivo function of alpha-carboxysomes and beta-carboxysomes in two model cyanobacteria. Plant Physiol *165*, 398-411.

Whitney, S.M., Houtz, R.L., and Alonso, H. (2011). Advancing our understanding and capacity to engineer nature's CO2-sequestering enzyme, Rubisco. Plant Physiol *155*, 27-35.

Woelfle, M.A., Ouyang, Y., Phanvijhitsiri, K., and Johnson, C.H. (2004). The adaptive value of circadian clocks: an experimental assessment in cyanobacteria. Current biology : CB *14*, 1481-1486.

Wollman, A.J., and Leake, M.C. (2015). Millisecond single-molecule localization microscopy combined with convolution analysis and automated image segmentation to determine protein concentrations in complexly structured, functional cells, one cell at a time. Faraday discussions *184*, 401-424.

Wollman, A.J., Shashkova, S., Hedlund, E.G., Friemann, R., Hohmann, S., and Leake, M.C. (2017). Transcription factor clusters regulate genes in eukaryotic cells. Elife *6*.

Wollman, A.J., Syeda, A.H., McGlynn, P., and Leake, M.C. (2016). Single-Molecule Observation of DNA Replication Repair Pathways in E. coli. Advances in experimental medicine and biology *915*, 5-16.

Wood, T.L., Bridwell-Rabb, J., Kim, Y.I., Gao, T., Chang, Y.G., LiWang, A., Barondeau, D.P., and Golden, S.S. (2010). The KaiA protein of the cyanobacterial circadian oscillator is modulated by a redox-active cofactor. Proc Natl Acad Sci U S A *107*, 5804-5809.

Woodger, F.J., Badger, M.R., and Price, G.D. (2003). Inorganic carbon limitation induces transcripts encoding components of the CO(2)-concentrating mechanism in Synechococcus sp. PCC7942 through a redox-independent pathway. Plant Physiol *133*, 2069-2080.

Woodger, F.J., Badger, M.R., and Price, G.D. (2005). Sensing of Inorganic Carbon Limitation in Synechococcus PCC7942 Is Correlated with the Size of the Internal Inorganic Carbon Pool and Involves Oxygen. Plant Physiology *139*, 1959-1969.

Wyka, T.P., and Luttge, U.E. (2003). Contribution of C3 carboxylation to the circadian rhythm of carbon dioxide uptake in a Crassulacean acid metabolism plant Kalanchoe daigremontiana. J Exp Bot *54*, 1471-1479.

Xu, M., Bernat, G., Singh, A., Mi, H., Rogner, M., Pakrasi, H.B., and Ogawa, T. (2008). Properties of mutants of Synechocystis sp. strain PCC 6803 lacking inorganic carbon sequestration systems. Plant & cell physiology *49*, 1672-1677.

Yang, Q., Pando, B.F., Dong, G., Golden, S.S., and van Oudenaarden, A. (2010). Circadian Gating of the Cell Cycle Revealed in Single Cyanobacterial Cells. Science *327*, 1522-1526.

Yeates, T.O., Crowley, C.S., and Tanaka, S. (2010). Bacterial microcompartment organelles: protein shell structure and evolution. Annual review of biophysics *39*, 185-205.

Yeates, T.O., Kerfeld, C.A., Heinhorst, S., Cannon, G.C., and Shively, J.M. (2008). Protein-based organelles in bacteria: carboxysomes and related microcompartments. Nat Rev Microbiol *6*, 681-691.

Yeates, T.O., Thompson, M.C., and Bobik, T.A. (2011). The protein shells of bacterial microcompartment organelles. Current opinion in structural biology *21*, 223-231.

Yokoo, R., Hood, R.D., and Savage, D.F. (2015). Live-cell imaging of cyanobacteria. Photosynth Res *126*, 33-46.

Zarzycki, J., Axen, S.D., Kinney, J.N., and Kerfeld, C.A. (2013). Cyanobacterial-based approaches to improving photosynthesis in plants. Journal of Experimental Botany *64*, 787-798.

Zarzycki, J., Erbilgin, O., and Kerfeld, C.A. (2015). Bioinformatic Characterization of Glycyl Radical Enzyme-Associated Bacterial Microcompartments. Applied and environmental microbiology *81*, 8315-8329.

Zarzycki, J., Sutter, M., Cortina, N.S., Erb, T.J., and Kerfeld, C.A. (2017). In Vitro Characterization and Concerted Function of Three Core Enzymes of a Glycyl Radical Enzyme - Associated Bacterial Microcompartment. Sci Rep *7*, 42757.

Zhang, M., Zhang, Z., He, K., Wu, J., Li, N., Zhao, R., Yuan, J., Xiao, H., Zhang, Y., and Fang, X. (2018). Quantitative Characterization of the Membrane Dynamics of Newly Delivered TGF-beta Receptors by Single-Molecule Imaging. Anal Chem.

Zhang, P., Eisenhut, M., Brandt, A.M., Carmel, D., Silen, H.M., Vass, I., Allahverdiyeva, Y., Salminen, T.A., and Aro, E.M. (2012). Operon flv4-flv2 provides cyanobacterial photosystem II with flexibility of electron transfer. The Plant cell *24*, 1952-1971.

Zhong, J., and Yan, J. (2016). Seeing is believing: atomic force microscopy imaging for nanomaterial research. RSC Advances *6*, 1103-1121.

Zhu, X.G., Long, S.P., and Ort, D.R. (2010). Improving photosynthetic efficiency for greater yield. Annual review of plant biology *61*, 235-261.

Appendix A

List of PCR primers used in this study. Gene primers are for amplification of gene fragments. F/R primers are for gene amplification, FP/KO primers are for redirections, SEG primers are for segregation screening.

Primers	Primer sequence （5’-3’）	Relative chapters
GFP 856-875 REV	TGGTGCAGATGAACTTCAGG	3
CcmK3_F	TCAAGCTGTTGGGCGAAGTC	4
CcmK3_R	AAGCATTGGCTTCGGCATCC	4
CcmK4_F	AGCGGTGATTGCAAGACACG	3,4
CcmK4_R	AACGCCAAAGCCACCATGAG	3,4
CcmK2_F	GCGGGAAGCTCAAGTTTGGG	4
CcmK2_R	ACTCGCTGGCCTATCCAGAC	4
CcmL_F	GAGCGAAGCTGTTACCTCAC	4
CcmL_R	GTTCCGTCCCAATCCGATAG	4
CcmM_F	ACGTGGACTCATGCCTATTG	4
CcmM_R	GGCAGAATAGTGCTGGAAAC	4
CcmN_F	CGAACACGAAGGTGAATACG	4
CcmN_R	AGTTCCGGTAGCTCCAACAG	4
CcmO_F	CGATGTCGCTGGACAGGTTC	4
CcmO_R	GTCGCGCTCGACTTGGATAC	4
CcmP_F	CTGTTTGCCACCGATGACAG	4
CcmP_R	CGCTCATGTTCAGCCAACTC	4
RbcL_F	CAACCTCGAAGCCGTGTTGC	3,4,5
RbcL_R	GGCTGTTAGCTCCGACTCTG	3,4,5
RbcS_F	ACGGGATTCACTTCCGTGTC	4
RbcS_R	GTGTTCCCTTGCCGTCGTAG	4
CcaA_F	AAACAGAGCTCGCCGTTATC	4
CcaA_R	GCATGAACGGGCATTCATTG	4
RbcX_F	AAGCAGGTTGGCAGCCTATC	4
RbcX_R	TCGCTGTCATCAAGGCATCG	4
CcmK3_FP_F	CCGATCGACTTCACGGCTGAATCCGAGCCCTTT CGGTCTCTGCCGGGCCCGGAGCTGCC	4

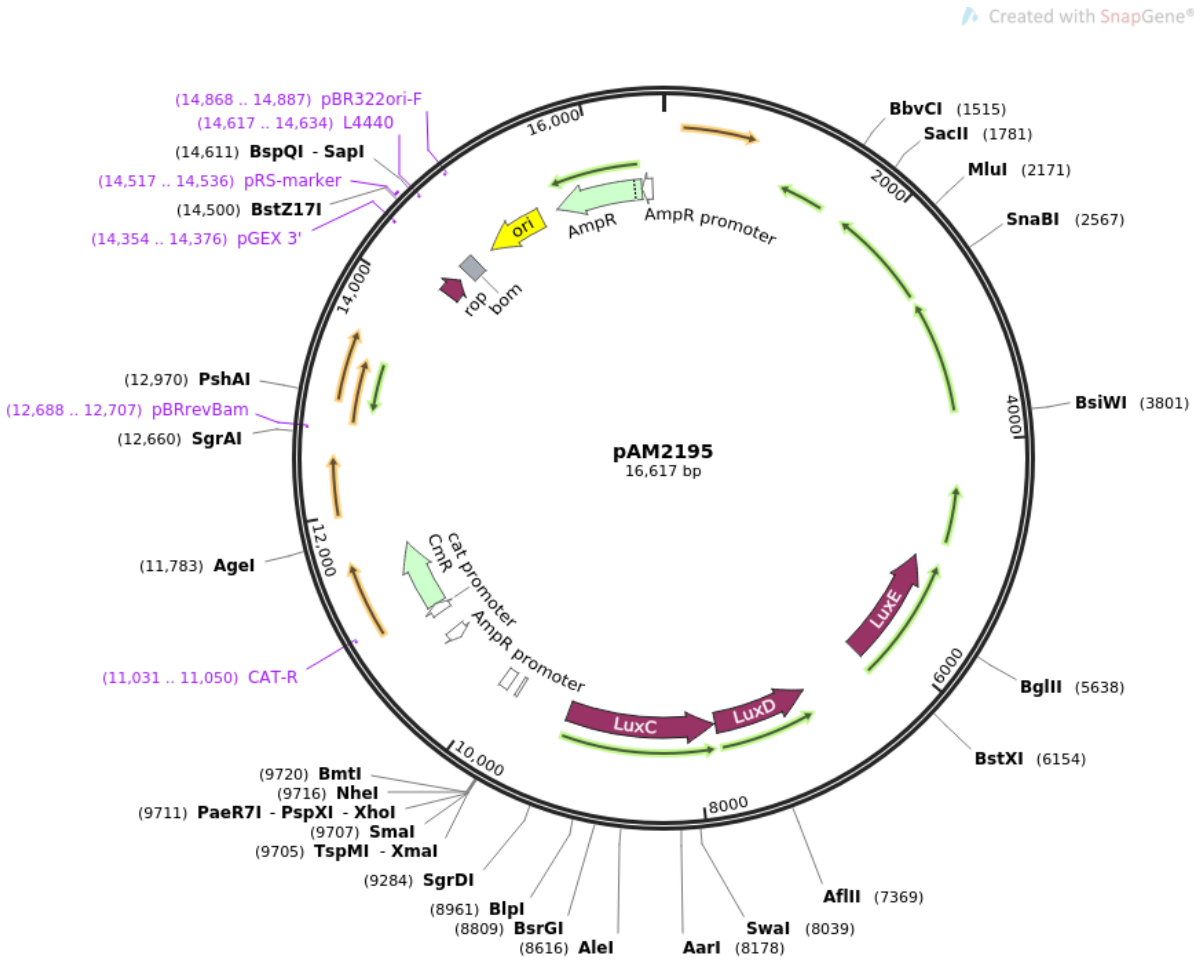
Primers	Primer sequence （5’-3’）	Relative chapters
CcmK3_FP_R	ACTCCTTCGCGTCGCAGTGTCGATTGTAAAGTC CGTGGCATTCCGGGGATCCGTCGACC	4
CcmK4_FP_F	TATCGACTCTCTGCCGAAGGAACTGGTAGTGGC CGCCGTCTGCCGGGCCCGGAGCTGCC	3,4
CcmK4_FP_R	CCCTCAGCCCAAATCAACCCTTTTAATCAGTCG CTACCCATTCCGGGGATCCGTCGACC	3,4
CcmK2_FP_F	CCGATTCGCTACACCGAAGCTGTTGAACAATTC CGCATGCTGCCGGGCCCGGAGCTGCC	4
CcmK2_FP_R	GATGAAAGGGCTATTGAAACGACTCAACAGCCC TCAACCATTCCGGGGATCCGTCGACC	4
CcmL_FP_F	GTGGAAAACCGCTCCGTCTACGACAAACGCGAG CACAGCCTGCCGGGCCCGGAGCTGCC	4
CcmL_FP_R	ATCCAGATCAATGACCAGCAGGGATTCAATCCC TGCCCAATTCCGGGGATCCGTCGACC	4
CcmM_FP_F	GCTCGTCGTCGCGTGGCTGAACTGTTGATTCAA AAGCCGCTGCCGGGCCCGGAGCTGCC	4
CcmM_FP_R	TGCATGAAGACAAGGGCGATCGCCGGATTCCCG GCACAGATTCCGGGGATCCGTCGACC	4
CcmN_FP_F	CAGTTTTTGCGGATGCGCCAGAGCATGTTCCCC GATCGCCTGCCGGGCCCGGAGCTGCC	4
CcmN_FP_R	ACAAAGTACCCTGCAGCTCCTAGAGCTGCTGTG CACATCATTCCGGGGATCCGTCGACC	4
CcmO_FP_F	GCCGAGCCCTTAGAGCTCCCCAATCCTCGTGAT GATCAGCTGCCGGGCCCGGAGCTGCC	4
CcmO_FP_R	TTTCAGCCAAAAACCAAGGGAATTACAGTGCTT TTTCCAATTCCGGGGATCCGTCGACC	4
CcmP_FP_F	CTGTCGGGACGTGAGCATCCCGGCGATCGCTCG CGGGAGCTGCCGGGCCCGGAGCTGCC	4
CcmP_FP_R	GCCGCAGAAATGCCAGATGTTCTGCATTCGCCA TGACCGATTCCGGGGATCCGTCGACC	4
RbcL_FP_F	AAAGAGATCAAGTTCGAATTCGAAACGATGGA CAAGCTCCTGCCGGGCCCGGAGCTGCC	3,4
RbcL_FP_R	GCAGCAACGCTCACTCCCCCAGCGATAGTCAGA GGCTCCATTCCGGGGATCCGTCGACC	3,4
GFP_internal_rev	TGGTGCAGATGAACTTCAGG	3
RbcS_FP_F	CAAACCGTGAGCTTCATCGTTCATCGTCCCGGC CGCTACCTGCCGGGCCCGGAGCTGCC	4
RbcS_FP_R	TGCCCAAAGGCAGGAGCAGCTATCAAGACAAA	4

Primers	Primer sequence （5’-3’）	Relative chapters
	TCAGGCTATTCCGGGGATCCGTCGACC	
CcaA_FP_F	CCTGAACAACAACAGCGGATTTATCGCGGCAATGCTAGCCTGCCGGGCCCGGAGCTGCC	4
CcaA_FP_R	ACACAGGCAGATCTCAGCAGGGTCGAAGATGCTTCGATCATTCCGGGGATCCGTCGACC	4
RbcX_FP_F	GCCACTCCTGATGATGCTTCTAATGCCTCCCATGCGGATCTGCCGGGCCCGGAGCTGCC	4
RbcX_FP_R	ATCGCGGCAGGCCCTTCAAAATCAACGTGTTGAACAATTATTCCGGGGATCCGTCGACC	4
CcmK3_insert_segF	ATCCCGAATCCACAAGAA	4
CcmK3_insert_segR	GCCCGCATGTAGCTCACAAT	4
CcmK4_insert_segF	GCCGAAGGAACTGGTAGTGG	3,4
CcmK4_insert_segR	GGGCTTCATCCCAGACATAG	3,4
CcmK2_insert_segF	GTTCTCCCGATTCGCTACA	4
CcmK2_insert_segR	AACAAATCTGCCAAGGTGC	4
CcmL_insert_segF	AAACCGCTCCGTCTACGAC	4
CcmL_insert_segR	TTGTTGGGCTCGGCATTCT	4
CcmM_insert_segF	GGATCGAGTTCGAGCGATGT	4
CcmM_insert_segR	TCACCACTGGCAAAGTAGCG	4
CcmN_insert_segF	CTAAGGTCTACGGCAAGGAA	4
CcmN_insert_segR	GCTGAGAATAGGCGGGAAG	4
CcmO_insert_segF	TGAGCCGCAGTCGGATAG	4
CcmO_insert_segR	GAAGGCGTTGACATCACTCT	4
CcmP_insert_segF	TCAGTGCGATTGGCAGTTT	4
CcmP_insert_segR	TCATTCCCGATCTTCGCCA	4
RbcL_insert_segF	CGTGAAGCTGGCAAGTGG	3,4
RbcL_insert_segR	GGAGGCAGGTACGAGAAAGT	3,4
RbcS_insert_segF	CAACATCAAGCAGTGCCAAAC	4
RbcS_insert_segR	CACCACCAATCACTCCAACG	4
CcaA_insert_segF	CCCTGACCGAAGAAGTGGC	4
CcaA_insert_segR	CAGGTAAGTCGGCAGGAAG	4
RbcX_insert_segF	GCCACTCCTGATGATGCTTC	4
RbcX_insert_segR	CACTTTGGGCGAGGGAGATT	4
KaiA_F	GATCGCAGACAAAGTGAAGG	5

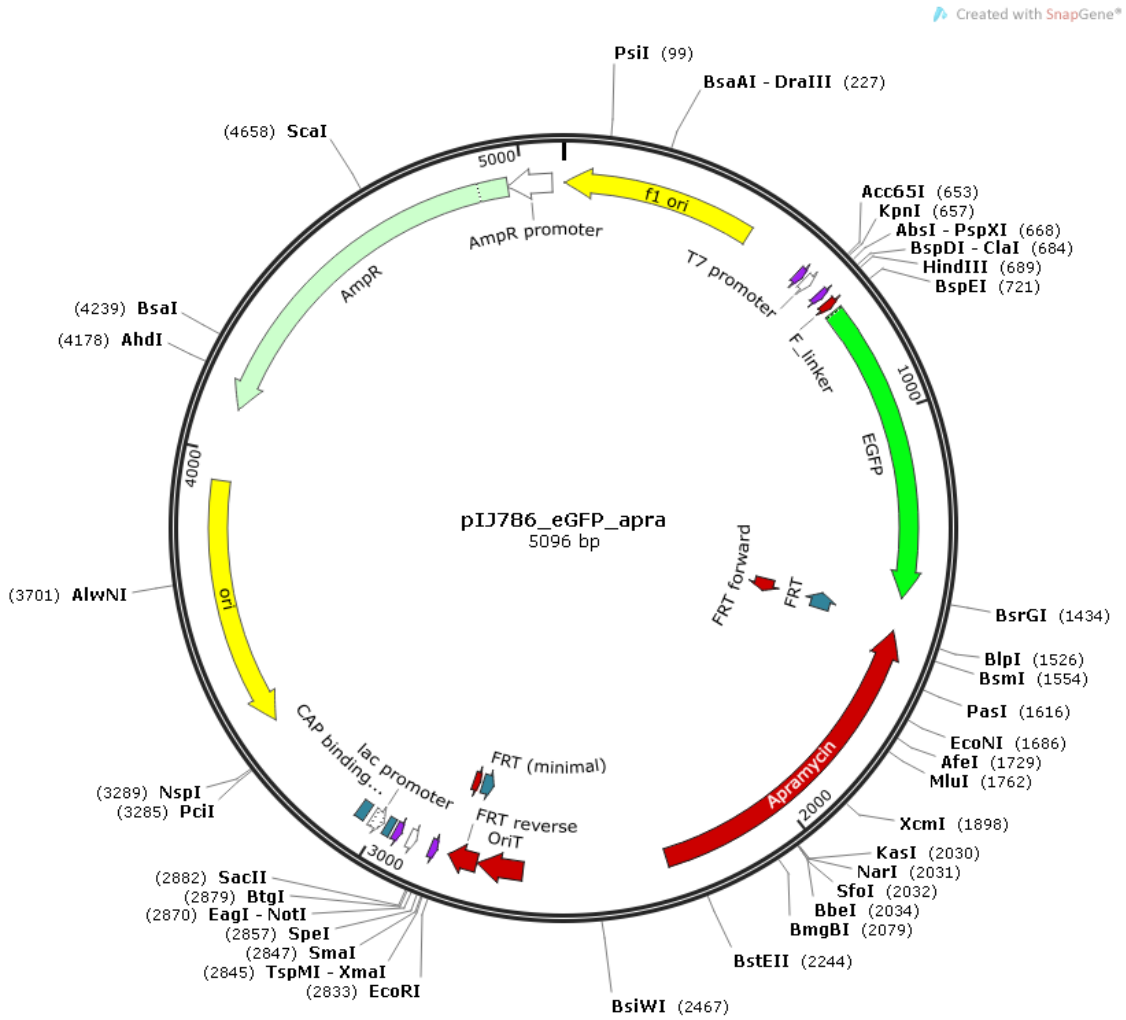
Primers	Primer sequence （5’-3’）	Relative chapters
KaiA_R	AAGAGGGTGAAGTCAGGTAG	5
KaiA_FP_F	CTTTGTGAGATGTATCGACGGTCTATCCCACGA GAAACCCTGCCGGGCCCGGAGCTGCC	5
KaiA_FP_R	GAGAGAAATTGAGCCGAGCTTAAGACCTCCTTT ACCTTTATTCCGGGGATCCGTCGACC	5
KaiA_KO_F	TCTGTCTGCAGACTCAGTCCTGACAGGAGCGAC TGCGTGATTCCGGGGATCCGTCGACC	5
KaiA_KO_R	AGAAATTGAGCCGAGCTTAAGACCTCCTTTACC TTTTCATGTAGGCTGGAGCTGCTTC	5
FKaiA KO SEG	ATGAGCTGCAGTGCTAGG	5
FKaiA FP SEG	CCGATGTTCCAGTCACCA	5
RKaiA FP/KO SEG	TTACGAGGGCTCATACGC	5

Appendix B

Plasmid map for pAM2195



Plasmid map for pIJ786_eGFP_apra



Plasmid map for pIJ786_eYFP_apra